Characterization of Protein-Metabolite and Protein-Substrate Interactions of Disease Genes

Dr. Alan Saghatelian Amanda Kathryn McFedries

Characterization of Protein-Metabolite and Protein-Substrate Interactions of Disease Genes

Abstract

Discovery of protein-metabolite and protein-substrate interactions that can specifically regulate genes involved in human biology is an important pursuit, as the study of such interactions can expand our understanding of human physiology and reveal novel therapeutic targets. The identification and characterization of these interactions can be approached from different perspectives. Chemists often use bioactive small molecules, such as natural products or synthetic compounds, as probes to identify therapeutically relevant protein targets. Biochemists and biologists often begin with a specific protein and seek to identify the endogenous ligands that bind to it. These interests have led to the development of methodology that relies heavily on synthetic and analytical chemistry to identify interactions, an approach that is complemented by *in vivo* strategies for validating the biological consequences of specific interactions.

Untargeted metabolomics is a powerful strategy for the identification of novel protein-small molecule interactions from various tissues that are relevant to target protein function during normal activity or within the context of a disease state. Here, this platform was used to identify the first endogenous ligands, prostaglandins A2, B2, D2 and E2, for the orphan nuclear receptor Nurr1, a gene that is also linked to Parkinson's Disease. These ligands function as inverse agonists of Nurr1 transcriptional activity and PGD2 and PGE2 have also been implicated in the neuroinflammatory pathway that contributes to PD progression.

Further demonstrating the value of these approaches in answering important chemical and biological questions, the work presented here on targeting the insulin degrading enzyme

(IDE) in the context of diabetes, illustrates the effectiveness of using small molecule tools to identifying novel endogenous substrates that play important roles in glucose homeostasis. Here the first potent and physiologically active IDE inhibitor was to discover that IDE regulates the abundance and signaling of glucagon and amylin, in addition to that of insulin. Importantly, under physiologic conditions that augment insulin and amylin levels acute IDE inhibition can lead to substantially improved glucose tolerance and slower gastric emptying. These findings transform our understanding of IDE's roles in glucose regulation, and demonstrate a potential therapeutic strategy of modulating IDE activity to treat type-2 diabetes.

Acknowledgements

I would like to thank Dr. Alan Saghatelian for all the support and guidance he has given me throughout the years. I am grateful for and appreciate all that I have learned while in his lab and I am especially thankful for Alan's ability to gather together such a bright and amazing group of labmates. Each member of the Saghatelian lab has helped me to succeed, in their own special way, and has made a positive impact on my life. This intelligent and diverse group of people made my graduate school experience rich with memory and exceedingly enjoyable. Thank you all for the love, support and guidance you have given me. I have learnt so much from each one of you.

This sentiment can easily be extended to all the members of Dr. David Liu's lab and to my classmates in the Molecular and Cellular Biology Department. Thank you all for the support, the fun, and the great scientific discussions. These years would have been pretty boring without Laila Akhmetova, Kyle McElroy, and Jamie Webster in particular.

I'd especially like to thank those with whom I've had the opportunity to collaborate with scientifically, especially Dr. Erin Adams, Dr. Yihong Wang, Dr. Sam Gellman, Dr. David Liu and Juan Pablo Maianti.

I would like to thank the faculty and the administration of the MCB department for being a consistent and dependable guiding force throughout all of my graduate school years. In particular Mike Lawrence, without him I'm sure the department would fall apart.

I'd like to thank my thesis committee, Dr. David Liu, Dr. Nathanael Gray and Dr. Andres Leschziner for their guidance and scientific discussions over the years, which have helped me to grow and succeed as a researcher.

Finally, I would like to thank my loving family for their support and generosity, especially my husband Tim Hawkins. Without you all I would not be the person that I am today.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACE	angiotensin converting-enzyme
cdiGMP	bis-(3’-5’)-cyclic dimeric guanosine monophosphate
CRABP2	cytosolic retinoic acid binding protein 2
CsA	cyclosporine A
CsnB	cytosporone B
CuACC	copper-catalyzed cycloaddition of azides and alkynes
DART	drug affinity responsive target stability
DBD	DNA binding domain
DGC	diguanylate cyclase
DIO	diet-induced obese
DPP4	dipeptidyl peptidase 4
DNA	deoxyribonucleic acid
DRaCALA	differential radial capillary action of ligand assay
DSF	differential scanning fluorimetry
DSLS	differential static light scattering
ERα	estrogen receptor alpha
ESI-TOF	electrospray ionization time of flight mass spectrometry
FABP2	fatty acid binding protein 2
FPLC	fast protein liquid chromatography
G6Pase	glucose-6-phosphatease
GCGR	G-protein coupled glucagon receptor
GLP-1	glucagon-like peptide 1
GLP-1	glucagon-like peptide 1
GPCR	G-protein coupled receptors
GST	glutathione s-transferase

GTT	glucose tolerance test
His6	polyhistidine
His6-Nurr1-LBD	His6-tagged-Nurr1LBD
IDE	insulin-degrading enzyme
IPGTTs	injected glucose tolerance tests
IPTG	isopropyl-B-D-1-thiogalactopyranoside
ITT	insulin tolerance tests
K_d	dissociation constant of the protein-ligand complex
LC-MS	liquid chromatography–mass spectrometry
MetE	S-methyltransferase
MMP1	matrix metalloprotease 1
MRM	multiple reaction monitoring
MudPIT	multidimensional protein identification technology
MW	molecular weight
MWCO	molecular weight cutoff
m/z	mass-to-charge ratio
nat-20(S)-yne	20(S)-hydroxycholesterol
NEP	neprilysin
NLN	neurolysin
NR	nuclear receptor
NR4A	nucleasr receptor subfamily 4A
OGTT	oral glucose tolerance tests
OD_{600}	optical density at 600 nm
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEPCK	phosphoenolpyruvate carboxykinase

PMI	protein-metabolite interactions
PSMI	protein-small molecule interactions
PTT	pyruvate tolerance test
ROS	reactive oxygen species
SEC	size exclusion chromatography
Shh	sonic hedgehog
SILAC	small molecule-affinity chromatography with stable isotope labeling of amino acids in cell culture
Smo	smoothened
SMTA	s-methyl thioacetimidate
SPROX	stability of proteins from rates of oxidation
StarD3	stAR-related lipid transfer domain protein 3
TBS	tris-buffered saline
T2D	type-2 diabetes
THOP	thimet oligopeptidase
Tm	melting temperature
UFA	unsaturated fatty acid
WT	wildtype

Chapter 1

Methods for the Elucidation of Protein-Small Molecule Interactions

This chapter was adapted from:

McFedries, A, Schwaid, A, Saghatelian, A. Methods for the Elucidation of Protein-Small Molecule Interactions. Chem Biol. 20, 667-73 (2013)

1.1 Introduction

A number of different types of molecular interactions enable life. These include the interactions between proteins, proteins and nucleic acids, and proteins and small molecules. Elucidating these interactions and understanding how they control biology is a major scientific goal. A number of approaches have been developed in recent years to identify protein-protein interactions [1-6] and protein-nucleic acid interactions [7-9]. Ribosome profiling, for example, elucidates interactions between the ribosome and RNA in cells to reveal novel sites of protein translation. While many methods exist for the characterization of biopolymer interactions, far fewer approaches exist to elucidate interactions between proteins and small molecules. In recent years, however, more of these methods are beginning to emerge.

The importance of protein-small molecule interactions (PSMIs) in drug discovery and protein-metabolite interactions (PMIs) in biology has driven the development of new methods that rely heavily on the integration of both synthetic chemistry and analytical chemistry. Here, we divide these methods into two categories: small molecule-to-protein and protein-to-small molecule strategies. This division separates problems that aim to identify the protein targets of a bioactive small molecule from problems focused on identifying the small molecule-binding partner of a suspected metabolite-binding protein.

1.2 Small-molecule affinity methods

One of the successes in using small molecules as affinity reagents is the identification of FKBP by the natural product FK506 [10]. This seminal work led to the discovery of new proteins and pathways that explained the mechanism of action of a potent class of immunosuppressant drugs [11]. In doing so, this work informed us about vital, but previously unknown, cellular pathways involved in the regulation of the cellular immune response. More generally, these studies highlighted the tremendous value of using bioactive small molecules to study biology. Other important examples, including the discovery of the histone deacetylase family with

trapoxin [12], reinforced the impact of chemistry in important biological discoveries, leading to the development of the field of chemical biology [13]. All of this is predicated on being able to use complex bioactive small molecules as affinity reagents, which often requires complex chemical syntheses and emphasizes the importance of organic chemistry in this problem.

The increased use of small molecule screening approaches in biology has led to the identification of many bioactive molecules, leading to an increased demand for methods that can elucidate PSMIs. Affinity-based methods are still the most common approach used and leading methods have learned to integrate these approaches with modern proteomics to accelerate targeted discovery. Ong and colleagues, for example, have combined small molecule-affinity chromatography with stable isotope labeling of amino acids in cell culture (SILAC) [14], a quantitative mass spectrometry (MS)-based proteomics strategy (using an ion trap MS), to identify PSMIs on a proteome-wide scale [14, 15] (Figure 1.1). To demonstrate the generality of this approach seven different compounds, including kinase inhibitors and immunophilin ligands, were studied in this first example. Derivatives of each of these compounds were prepared and linked to solid support by a carbamate linkage, affording small molecule derivatized beads.

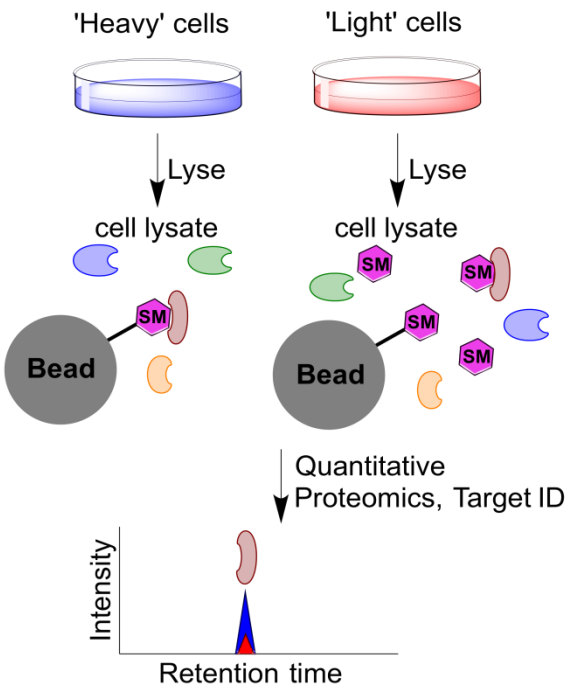


Figure 1.1 Affinity capture coupled to SILAC for small molecule target identification. Cells are grown in 'heavy' or 'light' media and subsequently lysed. These 'heavy' and 'light' cell lysates are separately incubated with beads that are modified with a small molecule of interest. Excess small molecule is added to the 'light' cell lysate and this soluble small molecule prevents any specific small molecule-protein interactions with the beads. After removal of the lysate bound proteins are eluted from the beads and the lysates are then combined for subsequent analysis using quantitative proteomics. The samples can be distinguished by mass spectrometry since proteins from each sample have different molecular weights, and therefore specific PMSIs can be identified by the higher concentrations of these 'heavy' proteins versus 'light' proteins.

In SILAC, cells are then grown in regular media (light) or specialized media (heavy) that replaces certain amino acids with stable isotope labeled derivatives (e.g. ^{13}C6-arginine and ^{13}C6,^{15}N2-lysine). The result is that the proteins in the 'heavy' cells have proteins that weigh

more than the exact same proteins in ‘light’ cells, and these proteins can be distinguished and quantified by mass spectrometry. SILAC is used to identify PSMIs by passing ‘light’ lysate over beads coated with the bioactive small molecule. Any proteins with affinity for the small molecule are retained. As a control, a soluble variant of the small molecule is added to the ‘heavy’ lysate before it is incubated with the small molecule-modified beads. This soluble compound has the effect of binding the target proteins and preventing their binding to the small molecules on the surface of the bead. The post-bead lysate samples are then combined and analyzed by mass spectrometry. The ratio of light-to-heavy can identify those proteins that are specifically enriched by the small molecule on the bead, and therefore identify any target proteins of the small molecule.

The results from these initial experiments were excellent. All the compounds used identified known PSMIs, and several revealed some novel interactions. Moreover, by using compounds with different affinities for their targets, this work demonstrated that this method can successfully identify PSMIs with affinities from the low nanomolar (26 nM) to micromolar (44 μM) range. This strategy has been applied to the identification of the primary targets of piperlongumine [16], a compound that was shown to selectively kill cancer cells in vitro and in vivo by targeting the stress response to reactive oxygen species (ROS) [17]. Overall, these types of affinity approaches have come to dominate the methods that are used to identify PMSIs. Examples include the identification of the nucleophosmin as a target of natural product avrainvillamide [18], the finding that cephalostatin A binds to specifically to members of the oxysterol binding-proteins, in the process revealing these proteins to be important in cancer cell proliferation[19].

Importantly, affinity methods are not limited to synthetic compounds or natural products, but can also be used with endogenous metabolites. Specifically, Nachtergaele and colleagues demonstrated a direct binding interaction between 20(S)-hydroxycholesterol and the oncoprotein smoothened (Smo), a key protein in the sonic hedgehog (Shh) pathway, using a

derivative of 20(S)-hydroxycholesterol (nat-20(S)-yne) that was immobilized onto a solid support using Sharpless' click chemistry [20]. This example highlights the generality of affinity-based approaches in identifying PSMIs. The only limitation appears to be the ability to access derivatives for immobilization by chemical synthesis and compounds that have or retain high affinity for their protein target. As scientists continue to gain interest in understanding the mechanism underlying bioactive small molecules, affinity-based methods will continue to be applied to many more target molecules.

1.3 Proteomic target identification

In cases where small molecules are difficult to modify, or the synthetic skill necessary to make such modifications are difficult to access, a new group of powerful proteomics methods to discover novel PSMIs can be used. These strategies rely on detecting differences in the stability between unbound and small-molecule bound proteins to identify the target(s) of a small molecule. Drug affinity responsive target stability (DARTS) is one such method [21].

DARTS relies on the fact that proteins are more stable when bound to a metabolite, which makes them less susceptible to proteolysis. Lysates are compared in the presence and absence of a small molecule. Proteins that bind to the small molecule are protected from proteolysis relative to the control sample, as indicated by mass spectrometry (Figure 1.2). Proof-of-concept experiments revealed that mammalian target of rapamycin (mTOR), for example, is less susceptible to proteolysis in the presence of E4, an mTOR kinase inhibitor. Moreover, DARTS is generally applicable and was used with other enzyme-inhibitor pairs, such as the COX2-celecoxib pair.

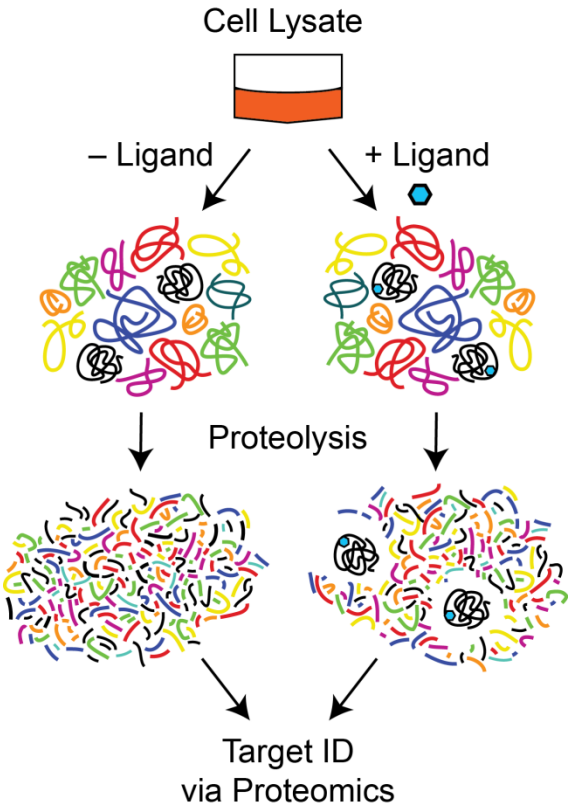


Figure 1.2 Drug Affinity Responsive Target Stability (DARTS). Aliquots of a protein lysate are mixed with either a small molecule (+ ligand) or solvent control (-ligand) to identify PSMIs. These samples are then subjected to limited proteolysis and compared by gel electrophoresis and quantitative mass spectrometry. Protein targets are identified as those proteins that display increased protease resistance in the presence of the small molecule.

Next, DARTS was used to characterize PSMIs for resveratrol, an anti-aging compound thought to act primarily through interactions with the sirtuin protein Sir1. Using a yeast and human lysates the authors discovered that eukaryotic initiation factor A (eIF4A) is a target of resveratrol, which demonstrates that DARTS is able to discover novel PSMIs. Importantly, the authors confirmed that at least some of the biology controlled by resveratrol is eIF4A dependent

because worms lacking eIF4A no longer show any anti-aging in the presence of resveratrol. Together these experiments demonstrate the value and utility of DARTS for discovering PSMIs.

Most recently, a proteomics method called Stability of Proteins from Rates of Oxidation (SPROX) was developed to identify PSMIs in complex cell lysates [22-25]. Rather than relying on non-specific proteolysis, which can make downstream mass spectrometry experiments difficult to interpret, SPROX relies on the irreversible oxidation of methionine residues by hydrogen peroxide to report on the thermodynamic stability of a protein's structure during chemical denaturation (Figure 1.3). Proof-of-concept SPROX experiments performed using yeast lysates validated this approach by identifying known binders of the immunosuppressant drug cyclosporine A (CsA) [23]. SPROX has also been used with resveratrol, identifying the known target, cytosolic aldehyde dehydrogenase, along with several novel interactions [25]. SPROX requires that target proteins contain methionine residues, and multidimensional protein identification technology (MudPIT) [26, 27] analysis of SPROX experiments using yeast lysates demonstrated that 33% of the detectable proteins contain a methionine.

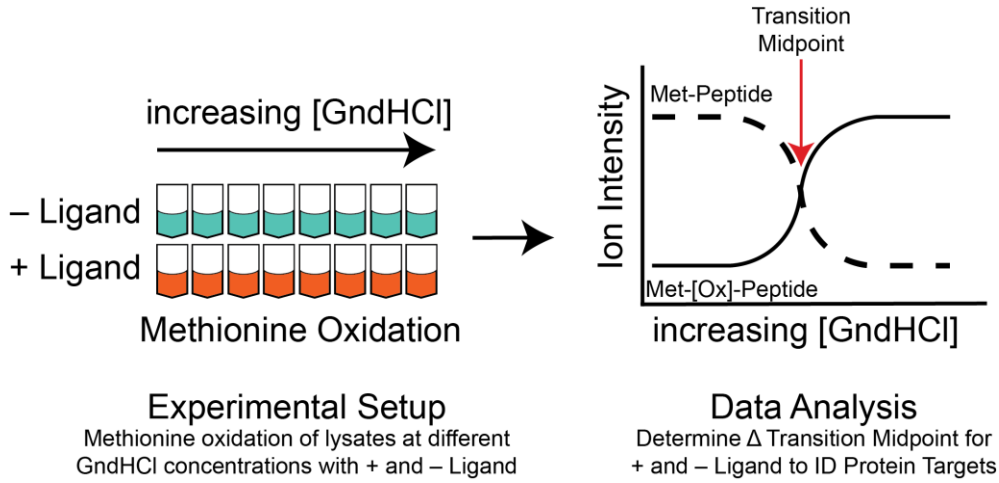


Figure 1.3 Stability of Proteins From Rates of Oxidation (SPROX). SPROX identifies the targets of a small molecule from a complex protein mixture by measuring the ligand-induced changes in the rate of methionine amino acid side chain oxidation by hydrogen peroxide. Aliquots of the cell lysate are incubated with either a small molecule or a solvent control and then incubated with increasing concentrations of guanidine hydrochloride.
Ligand induced difference in target protein unfolding will impact the rate of selective methionine oxidation by hydrogen peroxide. Quantitative proteomic is used to compare levels of reduced versus oxidized methionine-containing peptides in each sample set (small molecule versus control) to determine the rate of target protein oxidation as a function
of guanidine hydrochloride concentration. A shift of the transition midpoint for a protein between +ligand and –ligand samples indicates that the protein is a target for the small molecule in lysate.

A SPROX-like method that uses s-methyl thioacetimidate (SMTA) labeling to detect amidination differences of proteins and protein-ligand complexes during chemical denaturation has also been explored [28]. This method requires target proteins to contain lysine residues or have a buried N-terminus in the native state. This method is particularly useful when a ligand of interest is not stable in hydrogen peroxide and therefore cannot be investigated by SPROX.

SMTA labeling and SPROX complement each other, covering a wider range of the proteome [28]. Studying the thermodynamic stability of proteins under denaturing conditions in the presence and absence of ligand is an effective target protein identification strategy. However, the two described approaches require relatively large concentrations of ligand, in the uM to mM range, although they provide the flexibility to use a variety of downstream quantitative proteomic techniques.

1.4 Chemoproteomic target identification

Unfortunately, not all proteins are as active in lysates as they are within the context of a cell, and therefore some relevant PSMIs or PMIs may be missed when using lysates. Screening a spiroepoxide library for antiproliferative compounds, for example, revealed that the most relevant biological target of the active spiroepoxide is only targeted in a living cell, but was inactive once the cell was lysed [29]. Many molecular mechanisms can account for the difference between intact cells and lysates but the ultimate point is that it is difficult to exactly replicate cellular conditions in any type biochemical experiment. Since it is impossible to predict what PSMIs are sensitive to cellular conditions new chemoproteomic approaches have been developed to enable target identification within live cells.

These methods rely on the use of small molecules that can covalently label their protein targets so that intracellular labeling events can be detected after cell lysis. Manabe and colleagues demonstrated the value of this approach with a modified natural product derivative in an effort to identify its protein target [30]. Potassium isolespedezate, a metabolite known to induce nyctinastic leaf opening in the motor cells of plants belonging to the Cassia genus, was derivatized with an iodoacetamide for covalent crosslinking to its target and an azide to enable enrichment and identification of the isolespedezate target by conjugation to a flag peptide using click chemistry. This approach led to the identification of 5-methyltetrahydropteroyltriglutamate-homocysteine S-methyltransferase (MetE) as the isolespedezate target protein. This method,

while powerful, is mostly limited by the ability to synthesize natural product derivatives that can covalently label their target while maintaining the potency of the compound.

Most recently, a chemoproteomic strategy approach was developed for the identification of PMIs for the endogenous metabolite cholesterol. Cholesterol is a central metabolite with roles in membrane structure, metabolism, signaling and disease. While many important functions of this molecule are known, the full spectrum of proteins that interact with cholesterol is far from complete. Hulce and coworkers synthesized a series of cholesterol derivatives and controls containing a photocrosslinking diazirine group [31] (Figure 1.4). In practice, cells are irradiated by light after exposure to these sterol probes resulting in the covalent modification of any protein they bind. Addition of exogenous cholesterol blocks the overall labeling of these probes to validate that binding of these probes is occurring at cholesterol-specific binding sites. Subsequent to probe labeling, the probe is conjugated to biotin by copper-catalyzed cycloaddition of azides and alkynes (CuACC) [32] chemistry allowing for affinity enrichment of labeled proteins.

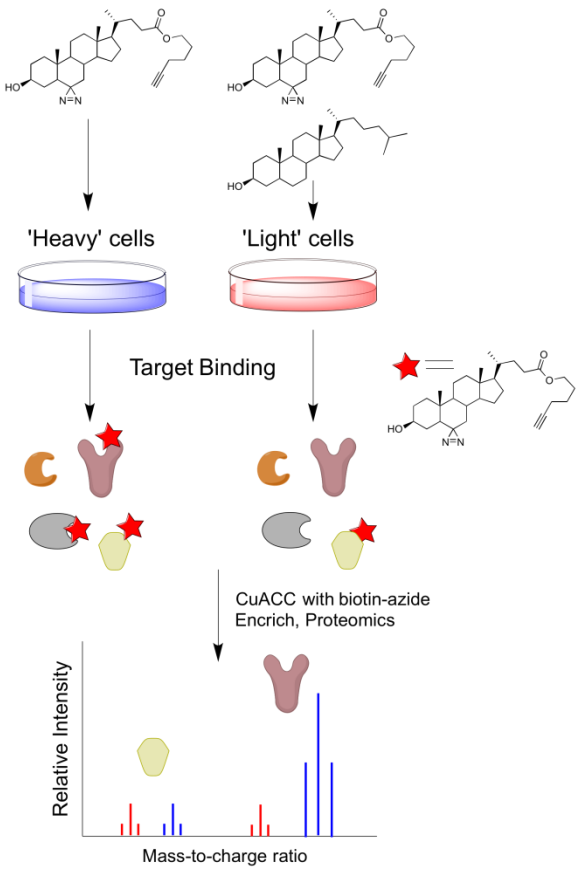


Figure 1.4 Identifying protein targets in cell culture. Cells grown in ‘heavy’ and ‘light’ media are treated with a cholesterol probe compound that has been modified to contain a diazirine moiety. In addition, excess cholesterol is also added to the ‘light’ sample and this will act to compete any specific cholesterol probe–protein interactions. The cholesterol probe is photocrosslinked to any bound protein targets in cells, and the cells are subsequently lysed. Any cholesterol probe–protein conjugates in this lysate are then modified with biotin using ‘click’ chemistry and labeled proteins are separated from cell lysate by affinity chromatography. The ‘heavy’ and ‘light’ fractions are then mixed and examined by quantitative proteomics. Proteins that specifically bind cholesterol will have a higher ratio of ‘heavy’ to ‘light’ in the mass spectrum.

Integrating these sterol probes with SILAC [14] enables the identification of sterol probe-target proteins. In these experiments, the probe is added to both 'heavy' and 'light' cells, but the 'light' cells also contain a competitor (cholesterol) to block binding. Subsequent analysis of the 'heavy' and 'light' samples identifies cholesterol-binding proteins as those proteins enriched in the heavy sample versus the control sample. The identification of several known cholesterol-binding proteins such as Scap, caveolin and HMG-CoA reductase validated the methodology. Subsequent analysis of the entire data set using various bioinformatics tools revealed that almost every major class of protein has members that bind cholesterol, including GPCRs, ion channels and enzymes. More broadly, the analysis also revealed an enrichment of proteins involved in neurological disorders, cardiovascular and metabolic disease, demonstrating the potential therapeutic insights that may eventually be provided by this data.

Together these examples demonstrate the utility of chemoproteomic approaches to identify PSMIs and PMIs, and highlight the power of these approaches to rapidly increase our understanding about the role of specific small molecules in biology.

1.5 Biophysical identification of small molecule binders

In many cases, the problem of identifying a PMI begins with interest in a particular protein. This protein may be a potential drug target or it may be suspected to require small molecule binding to regulate its activity. There are numerous cell-based assays for GPCRs [33-35] and nuclear receptors [36], for example. In general, these methods are highly effective. Such assays have already been reviewed extensively in the literature [33-36]. Instead, we've decided to focus on cell-free approaches that rely heavily on biochemical, biophysical and profiling methods to reveal PMIs for endogenous metabolites.

Biophysical screening methods provide an effective means for PMI discovery from endogenous metabolites. Differential scanning techniques were originally developed to optimize recombinant protein stability (i.e. melting temperature (Tm)) for purification and crystallography

[37]. Differential static light scattering (DSLS) and/or differential scanning fluorimetry (DSF) are the two most commonly used methods. DSLS measured denaturation by tracking temperature induced increases in the intensity of scattered light, while DSF measured increases in the fluorescence from the environmentally sensitive dye SYPRO Orange. Shifts in melting temperatures of > 2 °C were found to confidently represent binding events and conditions that enhanced protein stability, and thermal shifts > 4 °C increased the likelihood of positive results in crystallographic screens.

These methods have recently been extended to identify PMIs by measuring the effect of small molecules on the melting temperature (Tm) of proteins. The binding between a small molecule and protein stabilizes the protein structure (i.e. raises the Tm) (Figure 1.5). Recently, DeSantis and coworkers used DSF to identify natural estrogen receptor alpha (ERα) ligands from a library of molecules (not from a complex mixture) [38] . DSF successfully identified known natural ERα agonists, β-estradiol and estrone, demonstrating the utility of this assay in characterizing natural PMIs. The authors suggest that these assays will be useful in the identification of unknown nuclear receptor ligands in the future.

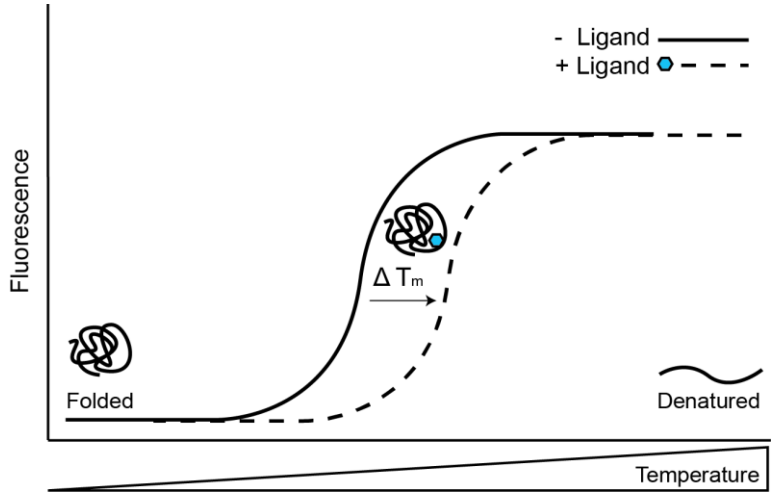


Figure 1.5 Thermostability Shift Assay. PSMI and PMIs can be identified in a high-throughput fashion by monitoring shifts in the melting temperature (Tm) of a protein-ligand complex versus a protein-solvent control. The environmentally sensitive dye SYPRO Orange emits when bound to hydrophobic amino acid residues and is used to monitor protein unfolding via differential scanning fluorimetry (DSF).

1.6 Affinity-based identification of small molecule binders

Alternatively, affinity based experiments using immobilized proteins are another option for the characterization of PMIs. In general, these assays provide the advantage that they can be performed with unmodified metabolite resulting in a reduced likelihood of false negatives. This approach relies on the fact that proteins can be immobilized without altering the secondary structure of the protein or interfering with ligand binding. The first examples of affinity methods relied on using radioisotopes to identify PMIs by measuring radioactivity of a protein after incubation and washing with a radiolabeled ligand [39, 40].

Most recently, this approach has been optimized in the form of a DRaCALA assay (Differential radial capillary action of ligand assay), which allows for the rapid high throughput identification of PMIs using radiolabeled metabolites [41]. DRaCALA utilizes the affinity of proteins for nitrocellulose membranes to sequester radiolabeled metabolites bound to protein. A solution containing the protein of interest and radiolabeled ligand is spotted on nitrocellulose. Unbound ligand will diffuse with the solvent throughout the membrane, whereas protein and ligand bound to protein will be immobilized at the point it was spotted. One advantage of DRaCALA is that it can avoid time-consuming protein purification by using whole cell lysate to identify PMIs as long as control cell lysate that does not express the metabolite-binding protein is available.

The one disadvantage of this method is that it requires foreknowledge as to what candidate metabolites should be tested, and in that each metabolite must be tested individually. Nevertheless, for certain cases DRaCALA provides a high-throughput means to identify ligand-binding proteins. The identification of prokaryotes that have proteins capable of binding bis-(3'-5')-cyclic dimeric guanosine monophosphate (cdiGMP), a metabolite important in biofilm formation, was accomplished by simply using lysates from 191 strains of P. aeruginosa and 82 other bacterial strains. The 'hits' in this assay corresponded to those bacteria that have diguanylate cyclase (DGC), as expected. This approach precludes the identification of unexpected PMIs or PMIs with novel metabolites. Still, DRaCALA remains a powerful approach for uncovering protein metabolite interactions.

The use of global mass spectrometry approaches allows for the scrutiny of a larger pool of metabolites, including novel metabolites, and can result in the unbiased identification of protein metabolite binding. Specifically, the use of global metabolite profiling enabled the development of a novel, unbiased, strategy for the identification of endogenous PMIs [42] (Figure 1.6). In this approach, recombinant proteins fused to an affinity tag—either GST or hexahistidine—are immobilized on a solid support. Incubation of these proteins with a

metabolite mixture, typically an extract containing the entire lipidome from a cell or tissue of interest results in the formation of a protein-metabolite complex on the bead. Following this incubation, the protein is washed, subsequently eluted from the solid support, and the eluent is then analyzed using a liquid chromatography–mass spectrometry (LC–MS) metabolite profiling platform (using an ESI-TOF for the MS). Quantitative comparison of the metabolite profiles between samples with and without protein reveals any metabolites that are enriched by the protein.

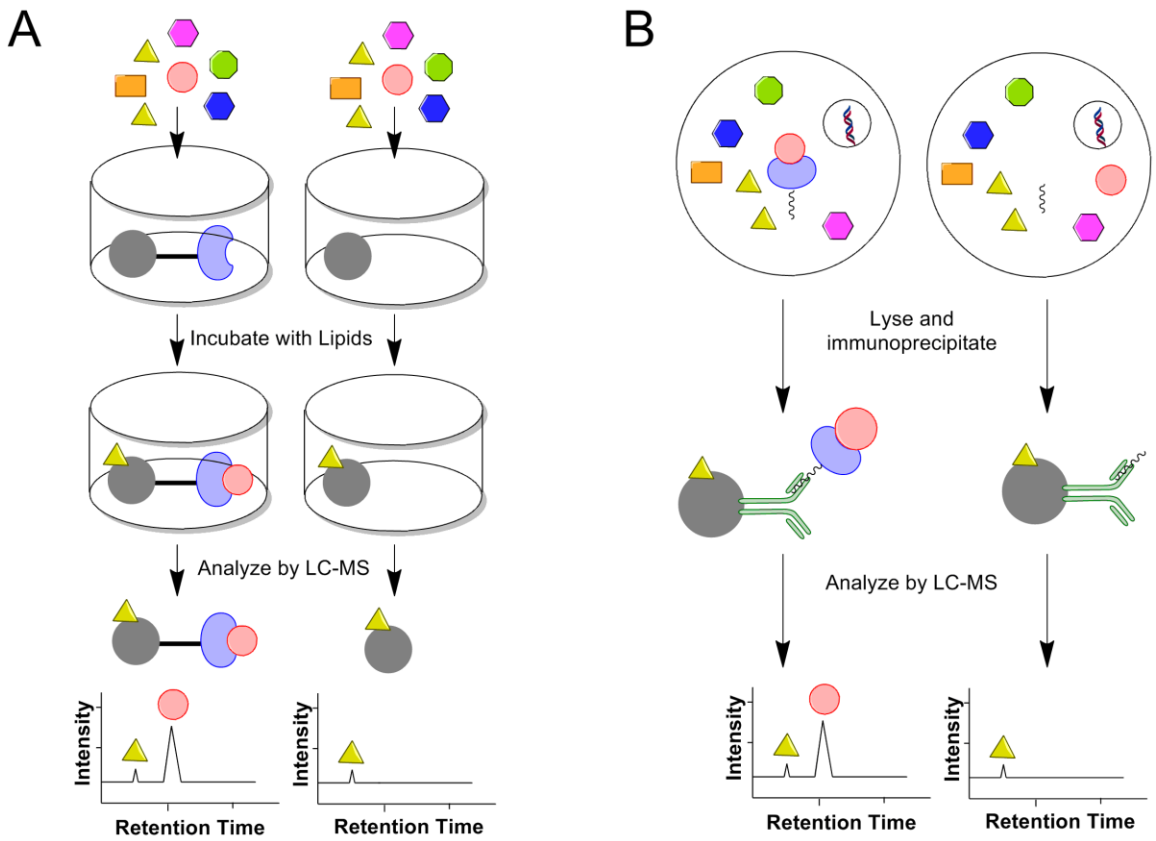


Figure 1.6. Affinity methods for elucidating PMIs. A) Protein (blue) is immobilized on a solid support (grey) and incubated with lipids from tissue lysate. The immobilized protein is then washed and the protein is eluted from the beads. The eluate is then analyzed by LC–MS and compared to eluate from a control sample (solid support with no protein) by in order to identify specific PMIs. B) A yeast strain bearing either the protein of interest fused to an IgG epitope tag, or the IgG epitope tag only, was lysed and immunoprecipitated using antibody labeled beads. Lipids were then extracted from the beads and examined by LC–MS. Comparison between the protein and control sample can be used to identify any specific PMIs.

This approach was developed using three different lipid-binding proteins with known ligands: cytosolic retinoic acid binding protein 2 (CRABP2), fatty acid binding protein 2 (FABP2) and StarD3. The strategy successfully identified specific PMIs for all three of these proteins, and

in doing so created a reliable method to pulldown protein metabolite interactions [42]. It does not require any explicit knowledge about the identity of the metabolite and can be conducted rapidly against a large pool of metabolites. Moreover, it was not susceptible to the enrichment of nonspecific metabolites. Although the examples presented here centered on lipids, there is no reason this approach could not be used for polar metabolites as well. This strategy has been successfully applied to other PMIs, including the discovery that arachidonic acid and docosahexanoic acid, polyunsaturated fatty acids, bind to the orphan nuclear receptor Nur77 [43].

Li and colleagues developed an exciting strategy to identify PMIs within yeast [44] (Figure 1.6). Specifically, they focused on interaction with yeast protein kinases as well as proteins that comprise ergosterol pathway. The proteins in this group were epitope tagged to enable their immunoprecipitation of from cellular lysates. These immunoprecipitated samples were analyzed by metabolomics to identify any endogenous metabolites bound to the proteins. As a control, each sample was also checked by SDS-PAGE to ensure pulldown of the target protein was successful.

Comparison of the metabolite profiles from these various samples revealed a number of new interactions. Many of the enzymes within the ergosterol pathway bound to sterol intermediates, which suggest that these molecules exert regulation on the pathway. Interestingly, it was also discovered that some of the proteins within the ergosterol pathway bind to pentaporphoryin, which was a complete surprise but helps explain a previous observation linking pentaporphoryin and ergosterol regulation. Furthermore, their analysis led to the discovery that a number of yeast protein kinases are regulated by ergosterol to highlight the generality of this method towards numerous protein classes. In aggregate, this data highlights the potential for unbiased PMI identification to greatly increase our current understanding of endogenous small molecule biology.

1.7 Future Directions

A successful library of approaches to determine PSMIs and PMIs has been developed, and is ready to be applied to identify unknown PMIs of interest. These approaches enable the discovery of novel interactions and are also designed to maximize the likelihood that the interactions are occurring in cells and tissues. The continued application of these methods will enrich our understanding of small molecule biology and also stimulate the development of improved methods for discovering these interactions. As demonstrated by the above examples this research area sits squarely at the interface of chemistry and biology and will greatly benefit from collaboration between future generations of chemists, biochemists and biologists.

1.8 References

1. Gingras, A.C., et al., *Analysis of protein complexes using mass spectrometry.* Nat Rev Mol Cell Biol, 2007. 8(8): p. 645-54.
2. Uetz, P.G., Loic; Cagney, Gerard; Mansfield, Traci A.; Judson, Richard S.; Knight, James R.; Lockshon, Daniel; Narayan, Vaibhav; Srinivasan, Maithreyan; Pochart, Pascale; Qureshi-Emili, Alia; Li, Ying; Godwin, Brian, Conover, Diana; Kalbfleisch, Theodore; Vijayadamodar, Govindan; Yang, Meijia; Johnston, Mark; Fields, Stanley; Rothberg, Jonthan M., *A comprehensive analysis of protein-protein interactions in Saccharomyces cerevisiae.* Nature, 2000. 403: p. 623-627.
3. Frei, A.P., et al., *Direct identification of ligand-receptor interactions on living cells and tissues.* Nat Biotechnol, 2012. 30(10): p. 997-1001.
4. Kaushansky, A., et al., *Quantifying protein-protein interactions in high throughput using protein domain microarrays.* Nat Protoc, 2010. 5(4): p. 773-90.
5. Hubner, N.C., et al., *Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions.* J Cell Biol, 2010. 189(4): p. 739-54.

6. Glatter, T., et al., *An integrated workflow for charting the human interaction proteome: insights into the PP2A system.* Mol Syst Biol, 2009. 5: p. 237.

7. Ingolia, N.T., et al., *The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments.* Nat Protoc, 2012. 7(8): p. 1534-50.

8. Johnson, D.S.M., Ali; Myers, Richard M.; Wold, Barbara, *Genome-Wide Mapping of in Vivo Protein-DNA Interactions.* Science, 2007. 316: p. 1497-1502.

9. Ingolia, N.T.G., Sina.; Newman, John R. S.; Weissman, Jonathan S., *Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling.* Science, 2009. 324: p. 218-223.

10. Harding, M.W., et al., *A receptor for the immuno-suppressant FK506 is a cis-trans peptidyl-prolyl isomerase.* Nature, 1989. 341(6244): p. 758-760.

11. Schreiber, S.L.C., Gerald R., *The Mechanism of Action of Cyclosporin A and FK506.* Immunology Today, 1992. 13(4): p. 136-142.

12. Taunton, J.C., Jon L.; Schreiber, Stuart L., *Synthesis of Natural and Modified Trapoxins, Useful Reagents for Exploring Histone Deacetylase Function.* Journal of the American Chemical Society, 1996. 118: p. 10412-10422.

13. Kijima, M., et al., *Trapoxin, an antitumor cyclic tetrapeptide, is an irreversible inhibitor of mammalian histone deacetylase.* Journal of Biological Chemistry, 1993. 268(30): p. 22429-35.

14. Ong, S.-E., et al., *Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics.* Molecular & cellular proteomics, 2002. 1(5): p. 376-386.

15. Ong, S.-E., et al., *Identifying the proteins to which small-molecule probes and drugs bind in cells.* 2009. 106(12): p. 4617-4622.

16. Raj, L., et al., *Selective killing of cancer cells by a small molecule targeting the stress response to ROS.* Nature, 2011. 475(7355): p. 231-4.

17. Adams, D.J.D., Mingji; Pellegrino, Giovanni; Wagner, Bridget K.; Stern, Andrew M.; Shamji, Alykhan F.; Schreiber, Stuart L., *Synthesis, cellular evaluation, and mechanism of action of piperlongumine analogs.* Proc. Natl. Acad Sci., 2012. 109(38): p. 15115-15120.

18. Wulff, J.E.S., Romain; Myers, Andrew G., *The Natural Product Avrainvillamide Binds to the Oncoprotein Nucleophosmin.* Journal of the American Chemical Society, 2007. 129(46): p. 14444-14451.

19. Burgett, A.W.G.P., Thomas B.; Wangkanont, Kittikhun; Anderson, D Ryan,; Kikuchi, Chikako; Shimada, Kousei; Okubo, Shuichi; Fortner, Kevin C; Mimaki, Yoshihiro; Kuroda, Minpei; Murphy, Jason P.; Schwalb, David J.; Petrella, Eugene C.; Cornella-Taracido, Ivan; Schirle, Markus; Tallarico, John A.; Shair, Matthew D., *Natural products reveal cancer cell dependence on oxysterol-binding proteins.* Nature Chemical Biology, 2011. 7: p. 639-647.

20. Nachtergaele, S., et al., *Oxysterols are allosteric activators of the oncoprotein Smoothened.* 2012. 8(2): p. 211-220.

21. Lomenick, B., et al., *Target identification using drug affinity responsive target stability (DARTS).* Proceedings of the National Academy of Sciences, 2009. 106(51): p. 21984-21989.

22. West, G.M., L. Tang, and M.C. Fitzgerald, *Thermodynamic Analysis of Protein Stability and Ligand Binding Using a Chemical Modification- and Mass Spectrometry-Based Strategy.* Analytical Chemistry, 2008. 80(11): p. 4175-4185.

23. West, G.M., et al., *Quantitative proteomics approach for identifying protein–drug interactions in complex mixtures using protein stability measurements.* Proceedings of the National Academy of Sciences, 2010. 107(20): p. 9078-9082.

24. Strickland, E.C., et al., *Thermodynamic analysis of protein-ligand binding interactions in complex biological mixtures using the stability of proteins from rates of oxidation.* Nat. Protocols, 2013. 8(1): p. 148-161.

25. DeArmond, P.D., et al., *Thermodynamic Analysis of Protein–Ligand Interactions in Complex Biological Mixtures using a Shotgun Proteomics Approach.* Journal of Proteome Research, 2011. 10(11): p. 4948-4958.

26. Washburn, M.P., D. Wolters, and J.R. Yates, 3rd, *Large-scale analysis of the yeast proteome by multidimensional protein identification technology.* Nature biotechnology, 2001. 19(3): p. 242-7.

27. Delahunty, C.M. and J.R. Yates, 3rd, *MudPIT: multidimensional protein identification technology.* BioTechniques, 2007. 43(5): p. 563, 565, 567 passim.

28. Xu, Y., et al., *Mass Spectrometry- and Lysine Amidination-Based Protocol for Thermodynamic Analysis of Protein Folding and Ligand Binding Interactions.* Analytical Chemistry, 2011. 83(9): p. 3555-3562.

29. Evans, M.J.S., Alan; Jorenson, Erik J.; Cravatt, Benjamin F., *Target discovery in small-molecule cell-based screens by in situ proteome reactivity profiling.* Nat Biotechnol, 2005. 23(10): p. 1303-1307.

30. Manabe, Y.M., Makoto; Ito, Satoko; Kato, Nobuki; Ueda, Minoru, *FLAG tagging by CuAAC and nanogram-scale purification of the target protein for a bioactive metabolite involved in circadian rhythmic leaf movement in Leguminosae.* Chemical Communications, 2010. 46: p. 469-471.

31. Hulce, J.J., et al., *Proteome-wide mapping of cholesterol-interacting proteins in mammalian cells.* Nat Methods, 2013. 10(3): p. 259-64.

32. Rostovtsev, V.V., et al., *A stepwise huisgen cycloaddition process: copper (I) ‐catalyzed regioselective "ligation" of azides and terminal alkynes.* Angewandte Chemie, 2002. 114(14): p. 2708-2711.

33. Civelli, O., *GPCR deorphanizations: the novel, the known and the unexpected transmitters.* Trends in pharmacological sciences, 2005. 26(1): p. 15-19.

34. Chung, S., T. Funakoshi, and O. Civelli, *Orphan GPCR research.* British journal of pharmacology, 2008. 153(S1): p. S339-S346.

35. Allen, J.A. and B.L. Roth, *Strategies to discover unexpected targets for drugs active at G protein-coupled receptors.* Annual review of pharmacology and toxicology, 2011. 51: p. 117-144.

36. Chawla, A., et al., *Nuclear receptors and lipid physiology: opening the X-files.* Science, 2001. 294(5548): p. 1866-1870.

37. Vedadi, M., et al., *Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination.* Proceedings of the National Academy of Sciences, 2006. 103(43): p. 15835-15840.

38. DeSantis, K., *Use of differential scanning fluorimetry as a high-throughput assay to identify nuclear receptor ligands* Nuclear Receptor Signaling, 2012. 10(e002).

39. Sundberg, S.A., *High-throughput and ultra-high-throughput screening: solution- and cell-based approaches.* Current Opinions in Biotechnology, 2000. 11: p. 47-53.

40. Bosworth, N.T., P.;, *Scintillation Proximity Assay.* Nature, 1989. 341: p. 167-168.

41. Roelofs, K.G.W., Jingxin; Sintim, Herman O.; Lee, Vincent T., *Differential radial capillary action of ligand assay for high-throughput detection of protein-metabolite interactions.* Proc. Natl. Acad Sci., 2011. 108(37): p. 15528-15533.

42. Tagore, R.T., Horatio R.; Homan, Edwin A.; Munawar, Ali; Saghatelian, Alan, *A Global Metabolite Profiling Approach to Identify Protein-Metabolite Interactions.* Journal of the American Chemical Society, 2008. 130: p. 14111-14113.

43. Vinayavekhin, N.S., Alan, *Discovery of a Protein-Metabolite Interaction between Unsaturated Fatty Acids and the Nuclear Receptor Nur77 Using a Metabolomics Approach.* Journal of the American Chemical Society, 2011. 133: p. 17168-17171.

44. Li, X., et al., *Extensive In Vivo Metabolite-Protein Interactions Revealed by Large-Scale Systematic Analyses.* 2010. 143(4): p. 639-650.

Chapter 2
Metabolomics Strategy Discovers Endogenous Inverse Agonists of the Orphan Nuclear Receptor Nurr1

2.1 Introduction

Nuclear receptors are a family of ligand-dependent transcription factors that respond to the environment of a cell by binding to a diverse battery of signaling molecules, such as oxysterols (LXR), vitamin D (VDR), estrogen (ER) and cortisol (GR)[1-3]. These receptors directly affect the expression of their target genes and play a central role in regulating genetic programs that control physiology related with numerous processes, including development and metabolism [2]. In addition to a role in normal physiology, nuclear receptor dysfunction is associated with diseases such as diabetes and cancer [1, 4]. As powerful regulators of physiology nuclear receptors are also targets of approved drugs, such as Advair (targeting GR for asthma and COPD) and Avandia (i.e. Rosiglitazone; targeting PPARγ for type II diabetes)[5] .

Nuclear receptors have a well conserved domain architecture composed of a highly variable N-terminal ligand-independent activation function (AF-1), a central DNA binding domain (DBD), a flexible hinge region and a C-terminal ligand-binding domain (LBD) that contains a ligand-dependent activation function (AF-2) [1, 6].The DBD contains two well conserved zinc fingers motifs that mediate receptor binding to specific DNA sequences, called hormone response elements (HRE), in the promoter region of target genes. The LBD is responsible for the recognition of the small molecule ligand, receptor dimerization, and the recruitment of co-regulator proteins to initiate transcription [2]. Members of the nuclear receptor superfamily share structural homology in their ligand binding domain (LBD), which is composed of 12 α-helices and 2-3 β-strands that are arranged as a three layered, antiparallel sandwich [1, 2].

At the molecular level the canonical nuclear receptor uses a conformational change upon binding of a small molecule ligand to control gene transcription. The mechanism of nuclear receptor activation can be described as a switch between associations with co-repressor and co-activator proteins driven by ligand-dependent positioning of the C-terminal alpha helix of the LDB (helix 12) [6]. In the absence of ligand, helix 12 occupies the co-activator binding surface and

facilitates co-repressor association with the receptor. Binding of an agonist induces a critical conformational change in the positioning of helix 12 that leads to the exchange of co-repressor (SMRT/NCOR and HDAC complexes) for co-activator proteins (HAT and chromatin remodeling complexes) [6-9] at a hydrophobic interface composed of residues from H12, H3 and H4 [1, 10].

Recently, the Saghatelian lab developed a metabolomics approach for elucidating protein-metabolite interactions and applied this approach to identify candidate endogenous ligands for the orphan nuclear receptor Nur77 (NR4A1). The NR4A subfamily contains two other members, Nurr1 (NR4A2) and Nor1 (NR4A3), also considered to be orphans or known members of the nuclear receptor family that lack any currently accepted ligand [11, 12]. We became interested in this class of nuclear receptors because they are widely expressed in energy demanding tissues (liver, brain, skeletal muscle and adipose), where they regulate essential pathways involved in metabolism and development such as gluconeogenesis (Nur77) [11, 13, 14], the development of dopaminergic neurons (Nurr1)[15, 16], and lipid, carbohydrate and energy metabolism (Nor1) [12].

The NR4A subfamily is characterized by atypical LBDs that lack both a canonical ligand binding pocket and a classical co-regulator interface [10, 17]. Previously, Nur77 and Nurr1 were reported to be constitutively active nuclear receptors that did not have or need a ligand to function. This conclusion was supported by the apo-LBD crystal structures of Nurr1 (PDB: 1OVL) [10] and Nur77 (PDB: 1YJE, 2QW4) [17] that indicated that the ligand binding domain of the receptors are filled with large, hydrophobic residues and, therefore, do not require a ligand. These crystal structures also showed that polar residues that did not associate with co-regulators populated the classical, hydrophobic, co-regulator interface. However, a novel co-regulator-binding cleft has been identified, composed of residues from helices 11 and 12. Co-repressor proteins are capable of interacting with this surface, but to date no co-activator proteins have been found to bind to the NR4A LBD [18, 19]. From this data it appears that the

N4RA family is unique in that it does not require a ligand to operate and does not use the canonical NR co-activators to modulate activity.

However, other data, at least in terms of ligand binding (Figure 1), suggest another story. First, prostaglandin A2 was identified as a NOR1 agonist through screening [20]. This indicates that there exist endogenous lipids that are able to bind and regulate the activity of members of the NR4A family. In addition, two small-molecule agonists of Nur77 have been reported, cytosporone B [21] and 1,1-bis(3-indolyl)-1-(p-methoxyphenyl)methane (BIM) [22], which indicate that Nur77 transcriptional activity can be regulated by the binding of a small molecule.

Regarding Nurr1, its activity can be activated by both benzimidadole (BEN) [23, 24] and isoxazolopyridinone (IXP) derivatives [24], but it is unclear if these compounds interact directly with the receptor. These finding prompted the search for a natural ligand for X using metabolomics. The application of this strategy revealed that Nur77 binds polyunsaturated fatty acids (PUFAs; arachidonic and docosahexanoic acids)[25]. Here, I apply this metabolomics approach to identify candidate ligands for the NR4A family member Nurr1 and characterize the impact of lipid binding on the receptor's transcriptional activity to identify bona fide endogenous ligands.

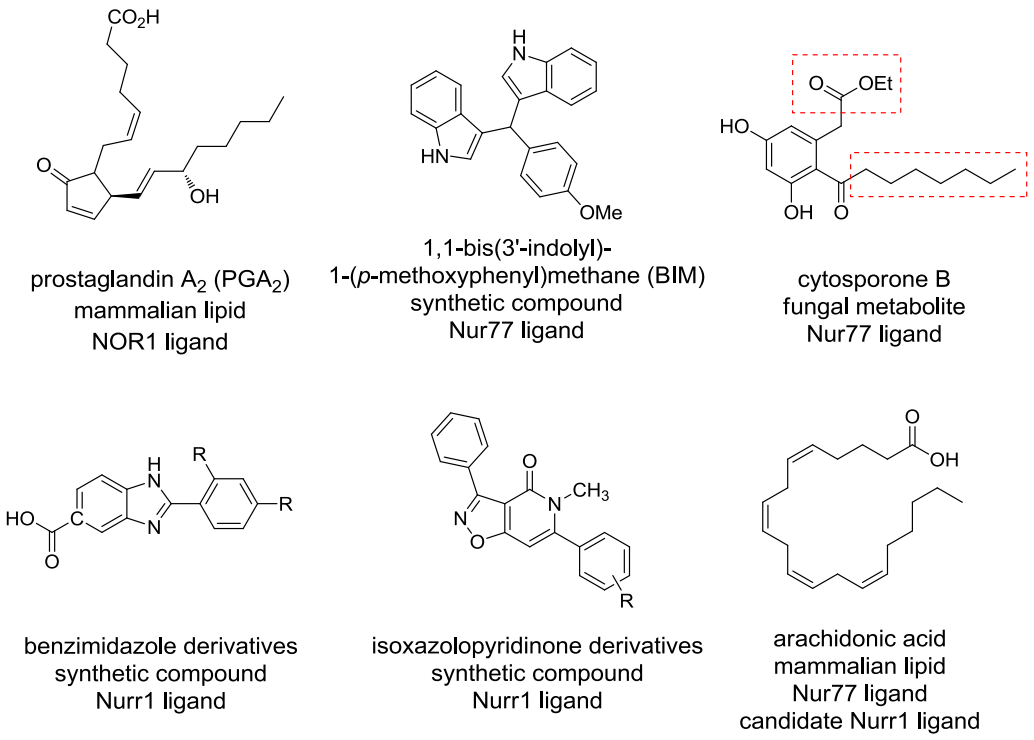


Figure 2.1 Ligands for the NR4A receptors include the Nor1 agonist prostaglandin A_2, the two Nur77 agonists BIM and Cytosporone B (boxed functional groups have properties reminiscent of endogenous lipids), and the Nurr1 activating derivatives of benzimidole and isoxazolopyridinone. Arachidonic acid has recently been identified inverse agonist of Nur77.

2.2 Identify natural ligands for Nurr1 through an untargeted metabolomics approach

To determine candidate ligands for the human Nurr1 receptor an untargeted metabolomics approach to identify protein-metabolite interactions (Figure 2) [26-28] will be used. Lipids for this experiments were prepared from mouse brain tissue because Nurr1 activity is essential for the development and maintenance of dopaminergic neurons in the brain.

The general strategy to identify protein-metabolite interactions begins with the recombinant expression and purification of the Nurr1 LBD (residues 328-598) with an N-terminal His_6 tag in *Escherichia coli* BL21(DE3) cells. To identify molecules that bind specifically to the

LBD it is immobilized on a metal ion affinity chromatography (IMAC) Sepharose 6 Fast Flow resin, where it is incubated with a pool of candidate lipids extracted from wildtype mouse brain. In this experiment, incubation of the resin in the absence of protein allows us to control for nonspecific binding of lipid to the solid support. After incubation the extract is filtered, the beads and protein are washed and the nuclear receptor is eluted from the resin. Eluted samples are then analyzed using an untargeted liquid-chromatography-mass spectrometry (LC-MS) based metabolomics platform that allows for the identification of lipids that are enriched in the protein containing sample versus the controls.

This system requires that use of internal standards to assist in the calculation of fold changes between metabolites in the sample sets [26-28]. Metabolite profiles, the LC-MS chromatograms from each sample, are compared using a software program called XCMS that aligns the corresponding peaks in each sample and calculates their intensities by integrating the area under the curves [29]. The differences in ion intensities between sample peaks are ranked by statistical significance (Student's t-test) and assigned a fold change value that is normalized to the internal standards [26-29]. Finally, we will identify the structure of enriched ions by using accurate mass (obtained during our standard MS experiments) and searching metabolite databases, such as METLIN, PubChem, or the lipid MAPS databases, for metabolites with similar molecular formulas [28].

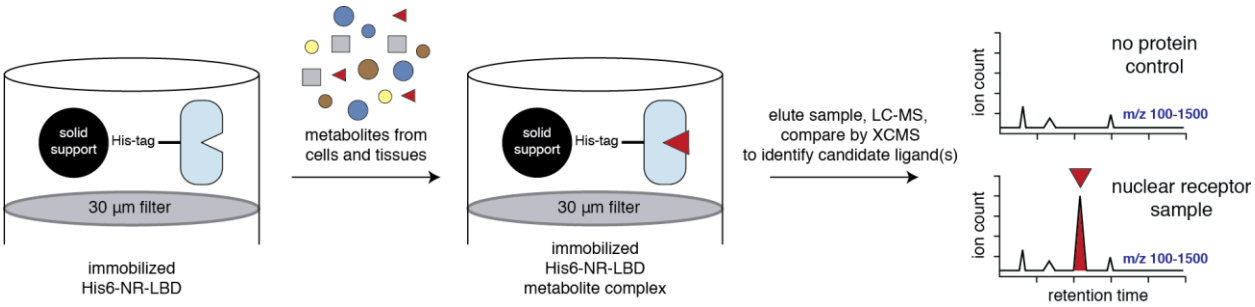


Figure 2.2 A liquid chromatography-mass spectrometry (LC-MS)-based metabolomics approach for characterizing endogenous nuclear receptor (NR) ligands. In the first step, metabolites isolated from cells and tissues are incubated with the ligand-binding domain (LBD) of an NR that has been immobilized onto a solid support (left). During the incubation any endogenous ligands can bind to the NR and form a protein-metabolite complex (middle). Unbound metabolites are then washed away and the nuclear receptor-ligand complex is eluted from the resin. NR-bound metabolites are identified by comparing the LC-MS chromatograms from the nuclear receptor sample to control sample using the program XCMS (right).

From the metabolomics assay I found that the Nurr1 LBD enriched polyunsaturated fatty acids (PUFAs) (p-value < 0.05 and fold-change > 2) in comparison to the no-protein resin control sample (Table 1). To be confident that we have identified endogenous ligands of Nurr1, the candidates must bind with physiologically relevant dissociation constants (nM to uM range)[30] to the receptor and be capable of regulating LBD dependent transcription. The secondary biochemical assays that were perform to characterize the lipid-protein interaction were circular dichroism (CD), which was used to measure conformational changes induced by lipid binding [25] (Figure 3) and an intrinsic tryptophan fluorescent assay, which was used to monitor saturation binding (Figure 5)[31]. Additionally, the intrinsic tryptophan fluorescent assay was used to screen prostaglandins (Figure 4) for their ability to bind the Nurr1 LBD as precedence for their ability to

regulate nuclear receptor function, particularly that of members of the NR4A family has already been established[20].

Enriched Lipids	Fold Change
Docosapentaenoic acid (DPA, C22:5)	10.1
Docosahexaenoic acid (DHA, C22:6)	7.7
Arachidonic Acid (AA, C20:4)	7.1
Oleic acid (C18:1)	7.1
Linoleic (C18:2)	5.3
No Enrichment (Controls)	**Fold Change**
Palmitic acid (PA, C16:0)	1.2
Stearic (C18:0)	1.2
Myristic (C14:0)	1.1

Table 2.1 Negative mode ions that are enriched by Nurr1-LBD greater than four-fold above the resin control across three separate experiments. Polyunsaturated fatty acids (PUFAs) are enriched, while saturated fatty acids are not. Positive mode analysis of samples did not reveal a consistently enriched set of ions all three data sets.

Quenching of intrinsic tryptophan (Trp) fluorescence can be used to measure candidate ligand binding. Excitation of Trp a 280nm will produce an emission at 330 nm when the residue is buried (hydropobic environment) and an emission at 350 nm when the residue is solvent accessible (polar environment). One caveat of this strategy is binding events with no impact on the local environmental of Trp residues will not be identified, allowing for the occurrence of false

positives. The Nurr1 LBD (2 uM) was incubated with candidate ligands (50 um) in a screen of saturated fatty acids, PUFAs and Prostaglandins (Figure 4). The fluorescence spectra is quenched upon treatment with the unsaturated fatty acids Oleic (18:1), Arachidonic (20:4) and Docosahexaenoic (22:6) as well as with the Prostaglandin B_2, while remaining relatively unchanged by the addition of saturated fatty acids and other prostaglandins. Furthermore, titration assays showed that the PUFAs and PGB2 were able to saturate binding (Figure 5). Indicating that the PUFAs and prostaglandins have the potential to be bona fide Nurr1 ligands.

To confirm the ability of the candidate ligands to modulate Nurr1 function, a cell based luciferase reporter assay was used to measure the regulation of downstream transcriptional activity in the presence of each candidate ligand. PUFAs and PGs did not have a significant impact on transcriptional activation on their own. However, treatment of cells with a benzimidazole activator (AROZ)[24] followed by treatment with the candidate ligands identified prostaglandins PGA2, PGB2, PGD2, and PGE2 as inverse agonists. PUFAs had no significant impact on Nurr1 LBD dependent transcriptional activity, but there could be a variety of reasons for this, such as their metabolism in these cell lines. The identification of PGs as inverse agonists of Nurr1 is a novel finding, and would suggested increased inflammation would reduce dopamine production, which is consistent with the implication of prostaglandin signaling in the development of Parkinson's Disease[32].

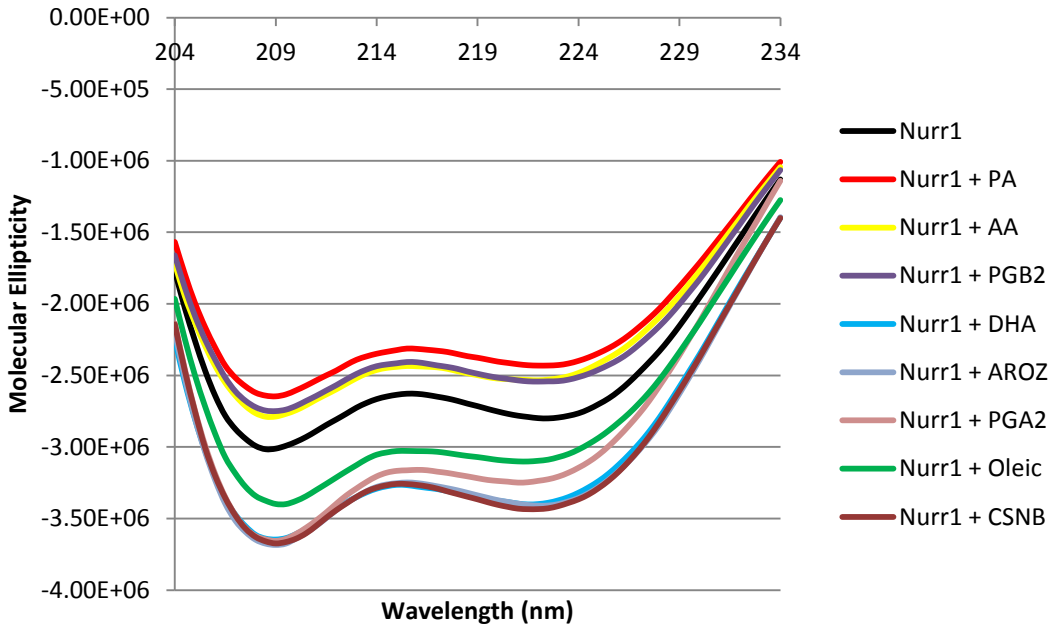


Figure 2.3 Circular Dichroism of Nurr1-LBD and candidate ligands

Circular Dichrosim spectra of Nurr1-LBD (5uM) after incubation with candidate ligands (50 uM) demonstrates that a variety of lipids can shift the secondary structure of the protein.

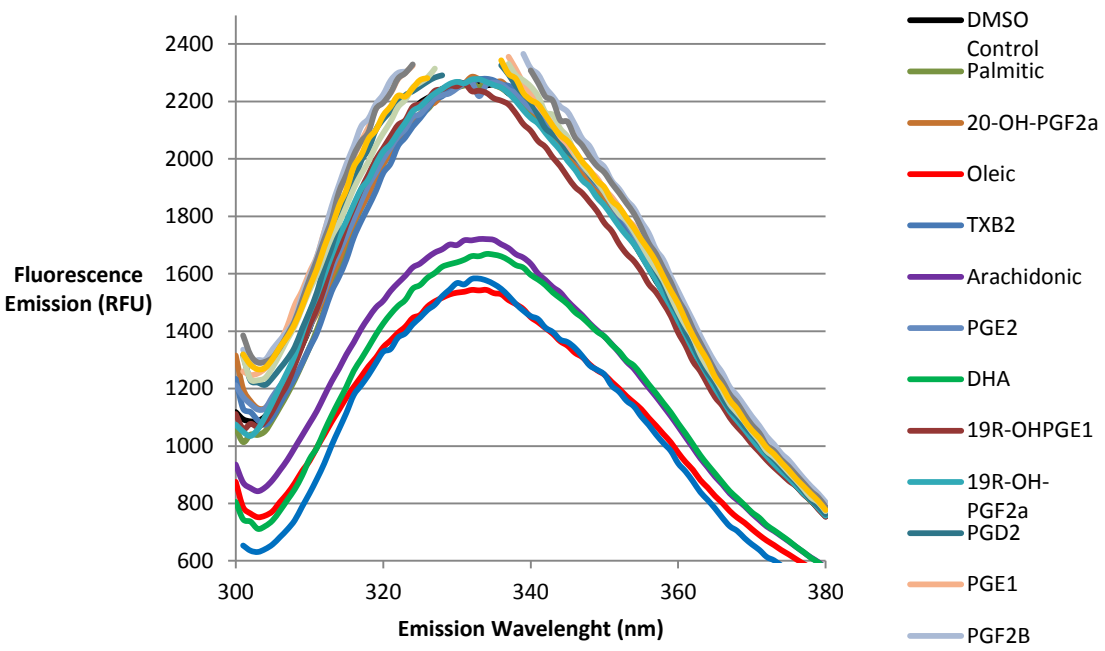


Figure 2.4 Intrinsic tryptophan fluorescence screen for lipid binding by Nurr1-LBD

The Nurr1-LBD was excited at 280 nM and its fluorescence monitored between 300-450 nM in the presence of 25 molar equivalents of candidate ligand. The fluorescence spectra is lowered upon treatment with the polyunsaturated fatty acids Oleic (18:1), Arachidonic (20:4) and Docosahexaenoic (22:6) as well as with the Prostaglandin B_2, while remaining relatively unchanged by the addition of saturated fatty acids and other prostaglandins.

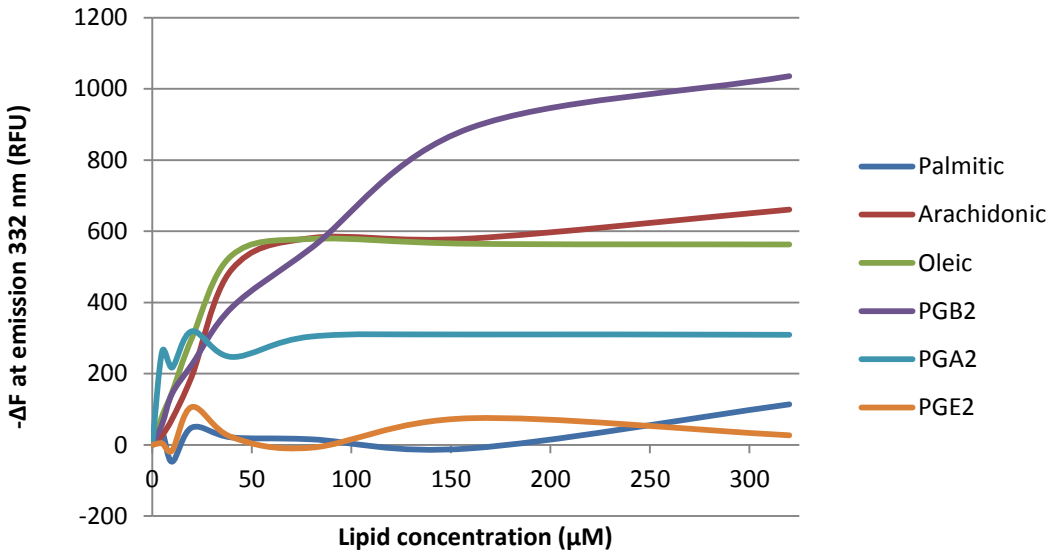


Figure 2.5 Saturation binding curves using intrinsic tryptophan fluorescence

Ligand-binding curves for palmitic acid (PA), arachidonic acid (AA), oleic acid, prostaglandin B_2 (PGB2), prostaglandin A_2 (PGA2) and prostaglandin E_2 (PGE2) were constructed by plotting the change in fluorescence (relative to 0μM lipid) at the maximum emission wavelength (332 nM) versus ligand concentration. With increasing concentrations of AA and Oleic acid (Kd ~33 μM) and PGB2 (Kd ~105 μM) you see a saturating fluorescence response, while there is relatively no change for PA of PGA2.

2.3 Conclusions

The PUFAs and PGs identified here are the first endogenous lipids identified to have a regulatory impact on Nurr1 function. This finding demonstrates the effectiveness and value of using an untargeted metabolomics approach to discover novel small molecule–protein interactions with bioactive properties. The fact that PGs, most specifically PGD2 and PGE2, have been implicated in the neuroinflammatory pathway contributing to the progression of Parkinson's Disease, positions this work to inform future research in targeting that pathway for therapeutic development. Additionally, future structural studies will be important for the identification and description of the NR4A unique mechanism of ligand-dependent

transactivation. This piece of work has deorphanized Nurr1 and will hopefully serve as a starting point for developing novel therapeutics that target the activity of this receptors for the treatment of Parkinson's disease.

2.4 Methods

Recombinant Nurr1LBD Expression (His6-Nurr1LBD) and Purification. Nurr1 LBD (residues 328-598) with an N-terminal His_6 tag was synthesized into the PJexpress411 vector by DNA2.0. 1-2L cultures of *Escherichia coli* BL21(DE3) cells that had been transformed with PJexpress411-His6-Nurr1LBD were grown with shaking in 2YT broth plus 50 µg/mL kanamycin until mid-log phase, then induced with 1mM IPTG for 4 hours. Cells were harvested by centrifugation (10 min at 10,000g) and stored at -80ºC until purification.

For purification of His6-Nurr1LBD protein, ≤2L cell pellets were resuspended in 35 mL lysis buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl, 20 mM imidazole, 1% TritonX-100) and were lysed by sonication at 4ºC. Lysate was cleared by centrifugation at 10,000g for 15 min at 4ºC. Soluble fraction was then loaded onto pre-equilibrated with lysis buffer nickel column (IMAC Sepharose 6 Fast Flow Resin from GE Heathcare). Column was washed with 50 mL lysis buffer and 100 mL was buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl, 20 mM imidazole), before application of 30 mL elution buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl, 500 mM imidazole). Eluate was concentrated using 10 kDa MWCO filter (Amicon). The protein was further purified by Superdex 200 10/300 GL column (GE Healthcare) in SEC buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl).

Tissue Extraction. An 8 mL solution of 2:1:1 CHCl3: MeOH: H2O was prepared for the extraction of 1 frozen mouse brain (Pel-freez Biologicals) in a 15-mL dounce tissue grinder. Sample was homogenized once ice until homogeneity was reached. Extract was then transferred to glass vials for centrifugation at 2,500 g for 20 minutes to separate organic and

aqueous layers. Organic (bottom) layer was removed and concentrated under stream of nitrogen gas. Extract was suspended in 200 uL of DMSO for use in enrichment experiments. Stored at -80C.

Ligand enrichment experiments with His6-Nurr77LBD. A 10-µL suspension of IMAC Sepharose 6 Fast Flow resin (GE Healthcare) was added to a 0.8 mL centrifuge column (Pierce Scientific #89868) and charged with Ni2+ ions. The resins were washed with water (500µL and then 2×100µL) and equilibrated equilibration with 3×200 µL of SEC buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl). Subsequently, the charged resins were incubated with either (i) 200 µL of 25µM His6-Nurr1LBD solution (experiment) or (ii) 200 µL protein SEC buffer (control). After 2 h incubation at 4 °C, the unbound protein was separated from the resins by centrifugation, and the resins were washed with 20 mM imidazole buffer (3× 200 µL). The lipid mixture (200 µL) was then prepared as a solution in protein buffer containing 4% (v/v) DMSO (i.e., the mixture of 8 µL lipid extracts in DMSO (or DMSO for no lipid control), and 192 µL protein buffer). After the resins were incubated with the lipid mixture for 1 h at room temperature, the lipid mixture was removed by centrifugation, and the resins were washed again for the last time with 20 mM imidazole buffer (3×200 µL). In the last step, the elution of the protein or any metabolite-bound protein was accomplished by incubating the resins with 100 µL of elution buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl, 500 mM imidazole) for 5 min prior to centrifugation. This step was repeated three times. The elution fractions were combined, and the sample was stored at -80 °C until it was analyzed by LC-MS.

LC-MS analysis of eluates from ligand enrichment experiment. LC-MS analysis was performed using an Agilent 6220 LC-ESI-TOF instrument. For the LC analysis in the negative mode, a Gemini (Phenomenex) C18 column (5 µm, 4.6 mm x 50 mm) and precolumn (C18, 3.5 µm, 2mm x 20mm) were used. Mobile phase A was a 95/5 water/methanol solution and mobile

phase B was a 60/35/5 isopropanol/methanol/water solution. Both mobile phase A and B were supplemented with 0.1% ammonium hydroxide as a solvent modifier. The flow rate was set to 0.5 mL/min throughout the duration of the gradient. The gradient began at 0% B for 12 minutes (first 10 minutes directed to waste) and then linearly increased to 100% B over the course of 40 min, followed by an isocratic gradient of 100% B for 6 min before equilibration to 0% B for 7min.

For LC analysis in the positive mode, a Luna (Phenomenex) C5 column (5 µm, 4.6 mm x 50 mm) and precolumn (C4, 3.5 µm, 2mm x 20mm) was used. Mobile phases A and B and gradient remained the same as used in negative mode, except 0.1% formic acid and 5 mM ammonium formate were used as solvent modifiers. The capillary voltage was set to 4.0 kV and fragmentor voltage to 100V. The drying gas temperature was 350ºC and flow rate was 10L/min, with the nebulizer pressure at 45 psi. Data was collected in both centroid and profile modes using a mass range of 100-1500 Da. Each run was performed using 80 µL injections of eluates from the ligand enrichment experiments and 45 µL for lipid mixtures.

LC-MS data analysis. Total ion chromatograms of eluates from the two sample sets (N=3) of His6-Nurr1LBD + lipids or no protein resin control +lipids was collected. Two step analyses was preformed, beginning with automated data analysis by XCMS followed by manual statistical analysis, confirmation and filtering of XCMS output files. To identify metabolites enriched by His6-Nurr1LBD the following parameters were used to filter XCMS output: fold changes ≥ 2, statistical signifigance ($p < 0.05$), and minimum MSII of 10,000 in samples of His6-Nurr1LBD + lipids. Visual inspection of EICs and elimination of isotopic peaks followed. Finally, enriched lipids were confirmed to be present in applied lipid mixture.

Circular Dichroism (CD) Spectroscopy. CD spectroscopy was performed on a Jasco J-710 spectropolarimeter. A chilled 200 uL mixture prepared containing 5 uM His6-Nurr1LBD (20 mM Tris-HCl pH 7.9, 100 mM NaCl) and 50 µM of candidate ligand in ethanol (or equivalent

volume of pure ethanol for control), with the ethanol concentration not to exceed 5% (v/v). The sample was then transferred to a quartz glass cuvette with 1-mm path length and scanned in quadruplicate at 20ºC across the wavelengths of 260 nm to 190 nm. To obtain the final CD spectrum, the raw data was subtracted i.e. protein containing sample minus buffer control, smoothed using Savitzky-Golay method with the convolution width of 5 and calculated for molar ellipticity values.

His6-Nurr1LBD intrinsic tryptophan fluorescence binding assay. In a 96 well standard, opaque plate 2 µM His6-Nurr1LBD protein was mixed with 50 µM candidate ligand (DMSO ≤5% v/v) in protein buffer (20 mM Tris-HCl pH 7.9, 100 mM NaCl). Data was collected on a SpectraMax M5 Microplate Reader using an excitation wavelength of 280 nm, with emission wavelength monitored between 300 to 450 nm, using a step size of 1 nm. Additional settings: PMT sensitivity set to high, no cutoff selected, autocalibrate turned on, automix turned off, and precision set to 14.

Transient Transfection and Reporter Analysis. To quantify Nurr1 transcriptional activity, CV1 cells or HEK293 cells in 48-well plates were transfected with 100ng/well NBRE-luciferase or NurRE-luciferase reporter, 50ng/well Nurr1, together with 50ng/well CMV-β-gal reporter (as internal control for transfection efficiency). NurRE-luc and Nurr1 plasmids were generously provided by Dr. Orla Conneely (Baylor College of Medicine). NBRE-luc plasmid was generously provided by Dr. David Mangelsdorf (UT Southwestern). All transfection was performed using FuGENE HD (Roche) (n=3) and repeated for at least three times. The cells were treated with each compound or vehicle control 24 hrs after transfection. Reporter assays were conducted 48 hrs after transfection, and luciferase activity was normalized by β-gal activity.

2.5 References

1. Huang, P., Chandra, V. & Rastinejad, F. Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics. *Annual Review of Physiology* 72, 247-272.
2. McEwan, I.J. Nuclear Receptors: One Big Family, in *The Nuclear Receptor Superfamily*, Vol. 505 3-18 (Humana Press, 2009).
3. Sonoda, J., Pei, L. & Evans, R.M. Nuclear receptors: Decoding metabolic disease. *FEBS Letters* 582, 2-9 (2008).
4. Chawla, A., Repa, J.J., Evans, R.M. & Mangelsdorf, D.J. Nuclear receptors and lipid physiology: opening the X-files. *Science* 294, 1866-1870 (2001).
5. Sladek, F.M. What are nuclear receptor ligands? *Molecular and Cellular Endocrinology* 334, 3-13 (2011).
6. Nagy, L. & Schwabe, J.W.R. Mechanism of the nuclear receptor molecular switch. *Trends in Biochemical Sciences* 29, 317-324 (2004).
7. Lonard, D.M., Lanz, R.B. & O'Malley, B.W. Nuclear receptor coregulators and human disease. *Endocr Rev* 28, 575-587 (2007).
8. Thakur, M.K. & Paramanik, V. Role of steroid hormone coregulators in health and disease. *Horm Res* 71, 194-200 (2009).
9. Rosenfeld, M.G., Lunyak, V.V. & Glass, C.K. Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes & Development* 20, 1405-1428 (2006).
10. Wang, Z. *et al.* Structure and function of Nurr1 identifies a class of ligand-independent nuclear receptors. *Nature* 423, 555-560 (2003).
11. Pearen, M.A. & Muscat, G.E.O. Minireview: Nuclear Hormone Receptor 4A Signaling: Implications for Metabolic Disease. *Mol Endocrinol* 24, 1891-1903 (2010).

12. Muscat, M.A.M.a.G.E.O. The NR4A subgroup: immediate early response genes with pleiotropic physiological roles. *Nuclear Receptor Signaling* 4 (2006).
13. Pei, L. *et al.* NR4A orphan nuclear receptors are transcriptional regulators of hepatic glucose metabolism. *Nat Med* 12, 1048-1055 (2006).
14. Chao, L.C. *et al.* Insulin resistance and altered systemic glucose metabolism in mice lacking Nur77. *Diabetes* 58, 2788-2796 (2009).
15. Le, W.D. *et al.* Mutations in NR4A2 associated with familial Parkinson disease. *Nat Genet* 33, 85-89 (2003).
16. Hawk, J.D. & Abel, T. The role of NR4A transcription factors in memory formation. *Brain Res Bull* 85, 21-29 (2011).
17. Flaig, R., Greschik, H., Peluso-Iltis, C. & Moras, D. Structural Basis for the Cell-specific Activities of the NGFI-B and the Nurr1 Ligand-binding Domain. *Journal of Biological Chemistry* 280, 19250-19258 (2005).
18. Codina, A. *et al.* Identification of a novel co-regulator interaction surface on the ligand binding domain of Nurr1 using NMR footprinting. *J Biol Chem* 279, 53338-53345 (2004).
19. Volakakis, N., Malewicz, M., Kadkhodai, B., Perlmann, T. & Benoit, G. Characterization of the Nurr1 ligand-binding domain co-activator interaction surface. *Journal of molecular endocrinology* 37, 317-326 (2006).
20. Kagaya, S. *et al.* Prostaglandin A(2) acts as a transactivator for NOR1 (NR4A3) within the nuclear receptor superfamily. *Biol Pharm Bull* 28, 1603-1607 (2005).
21. Zhan, Y. *et al.* Cytosporone B is an agonist for nuclear orphan receptor Nur77. *Nat Chem Biol* 4, 548-556 (2008).
22. Chintharlapalli, S. *et al.* Activation of Nur77 by Selected 1,1-Bis(3â€²-indolyl)-1-(p-substituted phenyl)methanes Induces Apoptosis through Nuclear Pathways. *Journal of Biological Chemistry* 280, 24903-24914 (2005).

23. Dubois, C., Hengerer, B. & Mattes, H. Identification of a potent agonist of the orphan nuclear receptor Nurr1. *ChemMedChem* 1, 955-958 (2006).

24. Hintermann, S. *et al.* Identification of a series of highly potent activators of the Nurr1 signaling pathway. *Bioorg Med Chem Lett* 17, 193-196 (2007).

25. Vinayavekhin, N. & Saghatelian, A. Discovery of a Protein–Metabolite Interaction between Unsaturated Fatty Acids and the Nuclear Receptor Nur77 Using a Metabolomics Approach. *J Am Chem Soc* 133, 17168-17171 (2011).

26. Tagore, R., Thomas, H.R., Homan, E.A., Munawar, A. & Saghatelian, A. A global metabolite profiling approach to identify protein-metabolite interactions. *J Am Chem Soc* 130, 14111-14113 (2008).

27. Kim, Y.G., Lou, A.C. & Saghatelian, A. A metabolomics strategy for detecting protein-metabolite interactions to identify natural nuclear receptor ligands. *Mol Biosyst* 7, 1046-1049 (2011).

28. Vinayavekhin, N. & Saghatelian, A. Regulation of alkyl-dihydrothiazole-carboxylates (ATCs) by iron and the pyochelin gene cluster in Pseudomonas aeruginosa. *ACS Chem Biol* 4, 617-623 (2009).

29. Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R. & Siuzdak, G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 78, 779-787 (2006).

30. Escriva, H., Delaunay, F. & Laudet, V. Ligand binding and nuclear receptor evolution. *BioEssays* 22, 717-727 (2000).

31. Petrescu, A.D., Hertz, R., Bar-Tana, J., Schroeder, F. & Kier, A.B. Ligand Specificity and Conformational Dependence of the Hepatic Nuclear Factor-4α (HNF-4α). *Journal of Biological Chemistry* 277, 23988-23999 (2002).

32. Lima, I.V.d.A., Bastos, L.F.S., Limborço-Filho, M., Fiebich, B.L. & de Oliveira, A.C.P. Role of Prostaglandins in Neuroinflammatory and Neurodegenerative Diseases. *Mediators of Inflammation* 2012, 13 (2012).

Chapter 3

Anti-Diabetic Activity of Insulin-Degrading Enzyme Inhibitors Mediated by Multiple Hormones

This chapter was adapted from:

Maianti, JP, McFedries, A, et. al.. Anti-Diabetic Activity of Insulin-Degrading Enzyme Inhibitors Mediated by Multiple Hormones. Nature. Accepted 2014.

3.1 Introduction

The dynamic interplay between the production and proteolytic degradation of peptide hormones is a key mechanism underlying the regulation of human metabolism. Inhibition of the peptidases and proteases that degrade these hormones can elevate their effective concentrations and augment signaling. In some cases, the resulting insights can lead to the development of novel therapeutics[1]. Dipeptidyl peptidase 4 (DPP4) inhibitors, for example, are anti-diabetic drugs that increase the concentration of the insulin-stimulating hormone glucagon-like peptide 1 (GLP-1), resulting in elevated insulin concentrations and lower blood glucose levels[2]. Researchers have speculated for at least six decades that insulin signaling could also be augmented to improve glucose tolerance by inhibiting its degradation[3,4]. The zinc metalloprotease insulin-degrading enzyme (IDE) is thought to be the primary enzyme responsible for inactivation of insulin in the liver, kidneys, and target tissues[3,4]. Recently, genome-wide association studies have identified predisposing and protective variants of the IDE locus linked to type-2 diabetes (T2D)[5-9], suggesting a functional connection between IDE and glucose regulation in humans.

Based on the known biochemistry of IDE, inhibition of this enzyme is expected to elevate insulin levels and augment the response to glucose[3,4]. While mice lacking a functional IDE gene ($IDE^{-/-}$ mice) have elevated insulin levels, counterintuitively these animals exhibit impaired, rather than improved, glucose tolerance[10,11]. Physiological studies with $IDE^{-/-}$ mice concluded that chronic elevation of insulin in these animals results in a compensatory lowering of insulin receptor expression levels, which leads to impaired glucose clearance following a glucose load[10,11]. This model raises the possibility that in the absence of such compensatory effects, acute inhibition of IDE may lead to improved physiological glucose tolerance.

Acute inhibition of enzymes is typically achieved through the use of small-molecule inhibitors, but the only previously reported IDE inhibitors are analogs of the linear peptide

mimetic Ii1 (Fig. 3.1)[12,13], which is highly unstable in vivo[11]. These compounds derive potency from a zinc-chelating hydroxamic acid group, which has the potential to interact strongly with other metalloproteases and metal-binding proteins[14]. These observations highlight the need for a selective small-molecule IDE inhibitor to characterize the biological functions and therapeutic relevance of this enzyme in vivo[15,16], uncoupled from confounding physiological adaptations that arise in IDE knock-out mice[10,11]. In this study we set out to discover potent and selective small-molecule IDE inhibitors that are active in vivo, and to use these compounds to reveal the physiological consequences of IDE inhibition.

3.2 Potent and selective IDE inhibitors

Ralph Kleiner and Juan Pablo Mainanti in the Liu lab led the discovery of small-molecule modulators of IDE, we performed in vitro selections on a previously described DNA-templated library of 13,824 synthetic macrocycles[17,18] for the ability to bind immobilized mouse IDE (Fig. A3.1). This library previously yielded highly selective Src kinase inhibitors,[19,20] and its unbiased design features and structural diversity suggested it may also contain ligands for other classes of proteins. Since each macrocycle in this library is attached to a unique DNA oligonucleotide that can be used to identify the structures of IDE-binding macrocycles[18,19], we used high-throughput DNA sequencing revealed the macrocycle-encoding DNA templates that were enriched following in vitro selection. Two independent IDE selection experiments, resulted in the identification of six putative IDE-binding macrocycles that share common structural features (Fig. 3a, Fig. A3.1).

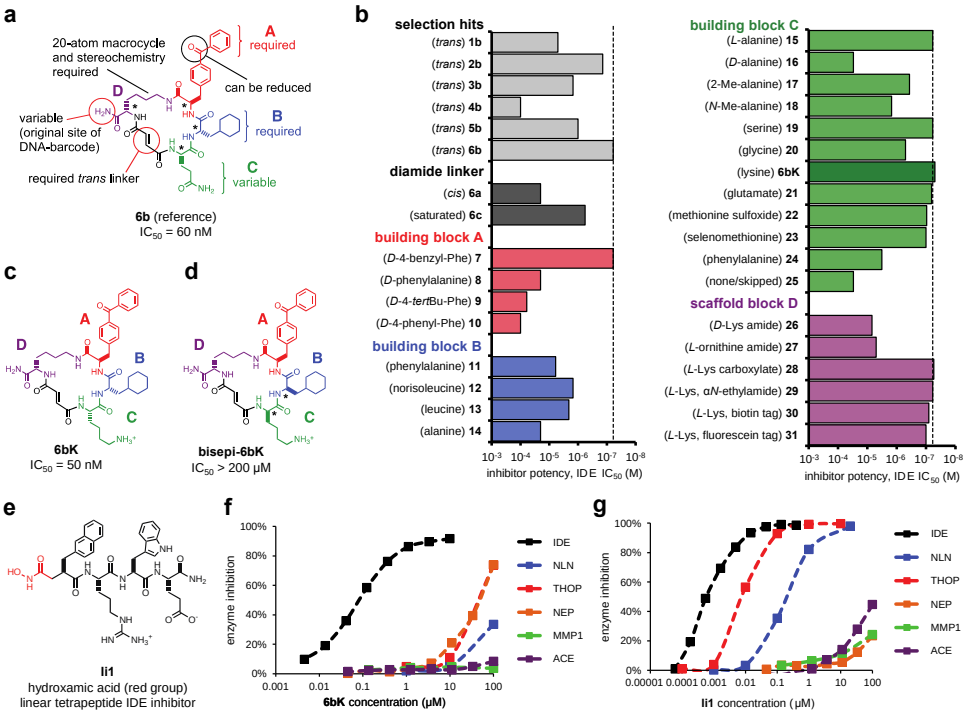


Figure 3.1 Discovery of potent and highly selective macrocyclic IDE inhibitors from the *in vitro* selection of a DNA-templated macrocycle library. **a**, Structure of the most potent hit from the IDE selection (**6b**) and a summary of the requirements for IDE inhibition revealed by the synthesis and evaluation of **6b** analogs. **b**, IDE inhibition potency of selection hits **1b** to **6b** and 30 structurally related analogs in which the linker, scaffold, and three building blocks were systematically varied. **c**, Structure of physiologically active IDE inhibitor **6bK**. **d**, Structure of the inactive diastereomer **bisepi-6bK**. **e**, Structure of the previously reported substrate-mimetic hydroxamic acid inhibitor **li1**[12]. (**F**) Selectivity analysis of macrocycle **6bK** reveals > 1,000-fold selectivity for IDE (□, IC_{50} = 50 nM) over all other metalloproteases tested. **g**, Inhibitor **li1**[12] inhibits IDE (■ IDE, IC_{50} = 0.6 nM), and also thimet oligopeptidase (■ THOP, IC_{50} = 6 nM) and neurolysin (■ NLN, IC_{50} = 185 nM), but not neprilysin (■ NEP), matrix metalloprotease 1 (■ MMP1), or angiotensin converting-enzyme (■ ACE). Human IDE shares 95 % primary sequence homology with mouse IDE[24], and both are inhibited by the macrocycles used in this study with similar potency (Fig. A2.2).

We synthesized these six macrocycles without their oligonucleotide templates and without the 5-atom macrocycle-DNA linker using solid- and solution-phase synthesis as either of two possible *cis*- or *trans*-alkene stereoisomers[18,19] (Fig. A3.1). Biochemical assays revealed that four of the six *trans*-macrocycles assayed were *bona fide* inhibitors of IDE with $IC_{50} \leq 1.5$ μM (Fig. A3.1). The most active inhibitor among the library members enriched in the selection, 20-membered macrocycle 6b (Fig. 3.1a), potently inhibited human IDE ($IC_{50} = 60$ nM). The ability of 6b to inhibit IDE proteolytic activity *in vitro* was confirmed by three complementary assays using a fluorogenic peptide, an insulin immunoassay, and a calcitonin-gene related peptide LC-MS assay (Fig. A3.2)[21].

We synthesized and biochemically assayed 30 analogs of 6b in which each building block was systematically varied to elucidate structure-activity relationships of these inhibitors (Fig. 3.1b). These studies revealed the structural requirements required for potent IDE inhibition by this new class of molecules, including a *trans*-fused 20-membered macrocycle, the stereochemistry of the macrocycle substituents, and the size, shape, and hydrophobicity of the A and B building blocks (Fig. 3.1a). In contrast to the strict requirements at positions A and B, different building blocks were tolerated at position C (Fig. 3.1b). Based on these results, we identified the inhibitor 6bK ($IC_{50} = 50$ nM, Fig. 3.1c) as an ideal candidate for *in vivo* studies because it exhibits enhanced water solubility relative to 6b, and can be readily synthesized on gram scale[16].

Because selectivity is a crucial feature of effective probes to elucidate physiological functions[16], we next characterized the protease inhibition selectivity of 6bK. This inhibitor exhibited ≥ 1,000-fold selectivity *in vitro* for IDE over all other metalloproteases tested: thimet oligopeptidase (THOP), neurolysin (NLN), neprilysin, matrix metalloprotease 1, and angiotensin converting-enzyme (Fig. 2.1f). In contrast, the previously reported substrate mimetic hydroxamic acid inhibitor Ii1[12] (Fig. 2.1e) is not as selective and it potently inhibits IDE ($IC_{50} = 0.6$ nM), THOP ($IC_{50} = 6$ nM), and NLN ($IC_{50} = 185$ nM) (Fig. 3.1g). The remarkable selectivity

of 6bK, in contrast with the known promiscuity of some active site-directed metalloprotease inhibitors[14], led us to speculate that the macrocycle engages a binding site distinct from the enzyme's catalytic site. This hypothesis was supported by double-inhibitor kinetic assays that revealed synergistic, rather than competitive, inhibition by macrocycle 6b and substrate mimetic Ii1 (Fig. A2.2).

3.3 Structural basis of IDE inhibition

In collaboration with Markus Selegur's lab, Juan Pablo and Zach Foda determined the molecular basis of IDE inhibition and selectivity by these macrocycles, we determined the X-ray crystal structure of catalytically inactive cysteine-free human IDE-E111Q[22] bound to 6b at 2.7 Å resolution (Fig. 3.2, Fig. A3.3a). The enzyme adopted a closed conformation and its structure is essentially identical to that of apo-IDE (PDB 3QZ2, RMSD = 0.257 Å). Macrocycle 6b occupies a binding pocket at the interface of IDE domains 1 and 2 (Fig. 3.2a), and is positioned more than 11 Å away from the zinc ion in the catalytic site (Figs. 3.2b and 3.2c). This distal binding site is a unique structural feature of IDE compared to related metalloproteases[23], and does not overlap with the binding site of the substrate-mimetic inhibitor Ii1[12]. The structure suggests that by engaging this distal site, the macrocycle competes with substrate binding (Fig. A3.3) and abrogates key interactions that are necessary to unfold peptide substrates for cleavage[24-26] (Fig. 3.2d).

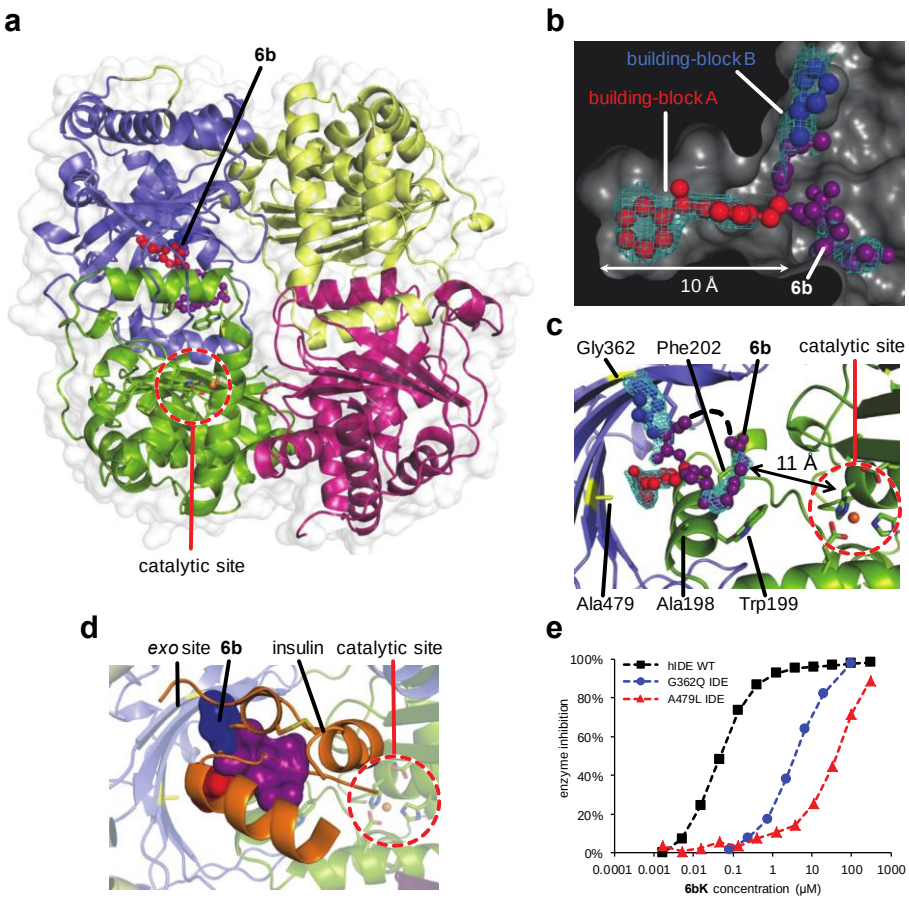


Figure 3.2 Structural basis of IDE inhibition by macrocycle 6b. **a**, X-ray co-crystal structure of IDE bound to macrocyclic inhibitor **6b** (2.7 Å resolution, pdb: 4LTE). IDE domains 1, 2, 3, and 4 are colored green, blue, yellow, and red, respectively. Macrocycle **6b** is represented as a ball-and-stick model, and the catalytic zinc atom is represented as an orange sphere. **b**, Electron density map (composite omit map contoured at 1σ) and model of IDE-bound macrocycle **6b** interacting with a 10 Å-deep hydrophobic pocket. **c**, Relative position of macrocycle **6b** bound 11 Å from the catalytic zinc atom. **d**, Overlay of the **6b** model (surface rendering) on the IDE:insulin co-crystal structure (pdb: 2WBY)[26]. **e**, Activity assays for wild-type or mutant human IDE variants in the presence of **6bK**. Mutagenesis of residues Ala479Leu (▲) and Gly362Gln (●) hindered the inhibition potency of **6bK** by > 600- and > 50-fold, respectively, compared to that of wild-type human IDE (■).

Because a section of the macrocycle was unresolved in the electronic density map, as observed with other ligands co-crystallized within the IDE cavity,[12] we sought to test the relevance of our structural model of the macrocycle:IDE complex (Fig. A3.3) to macrocycle:IDE binding in solution. We identified IDE mutations predicted by the structural model to impede 6b binding (Fig. 3.2e), prepared the corresponding mutant IDE proteins, and measured their abilities to be inhibited by 6b and 6bK. In addition, we also synthesized 6b analogs designed to complement these mutations and rescue inhibitor potency (Fig. A3.4). Building block A (*p*-benzoyl-phenylalanine in 6b) occupies a 10 Å-deep pocket in the crystal structure (Fig. 3.2b), defined by residues Leu201, Gly205, Tyr302, Thr316, and Ala479. As predicted by the structural model, mutation of Ala479 to leucine decreased the potency of inhibitors 6b and 6bK more than 600-fold, consistent with a significant steric clash in the binding site between Leu479 and the distal benzoyl group in building block A (Fig. 3.2e). Replacement of the *p*-benzoyl-phenylalanine building block with the smaller *tert*-butyl-phenylalanine, macrocycle 9, inhibited A479L-IDE with equal potency as wild-type IDE, consistent with the ability of the smaller macrocycle 9 to accommodate the added bulk of the leucine side chain (Fig. A3.4). Likewise, building block B (cyclohexylalanine in 6b) makes contacts in the structure with the peptide backbone of residues Val360, Gly361, and Gly362 located on the lateral β-strand 13 of IDE domain 2. These residues are thought to assist in unfolding of large peptide substrates by promoting cross-β-sheet interactions[24-26]. Mutation of Gly362 to glutamine decreased the inhibition potencies of 6b and 6bK at least 50-fold compared to wild-type IDE (Fig. 3.2e). A modified macrocycle (13) in which the cyclohexylalanine building block was replaced with a smaller leucine residue inhibited G362Q IDE and wild-type IDE comparably, consistent with a model in which the smaller B building block complemented the larger size of the glutamine side chain (Fig. A3.4).

Together, these structural and biochemical studies provide strong evidence for the proposed distal binding site of 6b and demonstrate the ability of the DNA-templated macrocycle

library to provide inhibitors that achieve unusual selectivity by targeting residues beyond the catalytic site. IDE is the only homolog of the M16 clade of metalloproteases, which is evolutionarily distinct from other zinc-dependent metalloprotease members and characterized by an inverted zinc-binding motif[23-26]. This distinct phylogenetic origin, together with the unusual binding mode of these macrocyclic inhibitors to IDE, may contribute to the unusual specificity of 6bK among proteases tested and encouraged us to explore the properties of 6bK *in vivo*.

3.4 Inhibition of IDE *in vivo*

Next we characterized the stability, physicochemical, and pharmacokinetic properties of 6bK. This macrocycle displayed significant stability during incubations with plasma and liver microsome preparations (74 % and 78 % remaining after 1 h, respectively), suggesting that this compound may be sufficiently stable *in vivo* to inhibit IDE activity. Plasma protein binding assays indicated that ~6 % of 6bK remains unbound and potentially available to engage its target at sites of insulin degradation, extracellularly or in early endosome compartments of target tissues[3,12]. To measure plasma half-life *in vivo*, we treated mice by intraperitoneal (i.p.) injection of 80 mg/kg 6bK using a formulation of sterile saline and Captisol, a β-cyclodextrin-based agent used to improve solubilization and delivery[27]. Plasma and tissue concentrations of 6bK were measured using isotope dilution mass spectrometry (IDMS) (Fig. A3.5). Plasma levels of 6bK were detectable 5 min post-injection, reached peak concentration (> 100 μM) at 60 min, and were maintained at a detectable level for at least 4 h (Fig. A3.5). This circulation time is within the timescale for standard physiological experiments with live animals[28,29]. We observed prompt biodistribution of 6bK into plasma, liver, kidney, and pancreatic tissues (Fig. A3.5). We did not detect any 6bK in the brain, suggesting that this macrocycle may not inhibit IDE within brain tissues, where IDE activity is needed for the clearance of β-amyloid peptides[10]. Indeed, brain tissue levels of Aβ(40) and Aβ(42) peptides in mice injected with 6bK were unchanged 2 h-post injection compared to vehicle-treated controls (Fig. A3.6), consistent with

the inability of 6bK to inhibit IDE in the brain. Collectively, these observations suggested the viability of 6bK as a potential *in vivo* IDE inhibitor probe in peripheral tissues of mice[16].

To evaluate the ability of 6bK to inhibit IDE activity *in vivo*, we subjected non-fasted mice to insulin tolerance tests (ITT)[29] following a single injection with 6bK (80 mg/kg) formulated in Captisol[27]. The insulin injections were carried out 30 min post-injection, corresponding to the time of highest 6bK concentration in plasma (approximately 100 μM, ~1000-fold the IC_{50} for mouse IDE). Following a subcutaneous insulin injection, mice treated with 6bK experienced lower hypoglycemia and higher insulin levels compared to vehicle controls ($p < 0.01$, Fig. A3.6 and Fig. 3.4b). In 1955, Mirsky used a similar ITT assay to suggest the feasibility of insulin stabilization by injecting rats with preparations of an undefined endogenous IDE inhibitor crudely fractionated from bovine livers (presumably a competitive substrate; see Appendix Table 1)[30].

Our experiments with 6bK provide the first evidence that a well-defined, selective, and physiologically stable pharmacological IDE inhibitor can augment the abundance and activity of insulin *in vivo*. Testing the macrocycle at several doses established an effective dose of 6bK of ~2 mg/mouse i.p., representing 80 mg/kg for lean C57BL/6J mice (25 g), and 60 mg/kg for diet-induced obese (DIO) mice (35-45 g). These inhibitor doses were well tolerated, did not produce adverse reactions, or body weight loss (Fig. A3.6), and did not induce detectable behavioral abnormalities.

3.5 Anti-diabetic activity of IDE inhibition during oral glucose administration

To determine the physiological consequences of acute IDE inhibition *in vivo*, we evaluated glucose tolerance of mice treated with 6bK. We used two methods of glucose delivery, either oral gavage or i.p. injection,[28] and two different mouse models, lean or DIO mice[31,32]. These four conditions were chosen to survey the role of IDE activity under a broad range of endogenous insulin levels and insulin sensitivity[28,31]. Oral glucose administration, for example, results in greater insulin secretion compared to injected glucose delivery (Fig. A3.7). Passage

of glucose through the gut causes the release of GLP-1, which strongly augments glucose-dependent insulin secretion[2,31]. This phenomenon is referred to as the 'incretin effect' (Fig. A3.7) and is magnified in DIO mice[31]. In addition, DIO mice display hyperinsulinemia and insulin resistance compared to lean mice, enabling us to test the consequences of IDE inhibition in a model that resembles early type-2 diabetes in humans[32].

In all glucose tolerance experiments we included two control groups: vehicle alone, and the inactive isomer[16] bisepi-6bK (Fig. 3.1c), which is identical to 6bK in chemical composition and bond connectivity, but has virtually no IDE inhibition activity (IC_{50} > 100 μM, Fig. 3.1c, see also Fig. A2.8). As expected, administration of 6bK alone, without a glucose challenge, did not significantly alter basal blood glucose or hormone levels compared to control treatments (Fig. A3.8). We then examined the effect of 6bK on blood glucose levels during an oral glucose tolerance test (OGTT). Lean or DIO mice were fasted overnight for these experiments, and then treated with a single dose of 6bK, vehicle alone, or a matching dose of inactive bisepi-6bK. After 30 minutes, glucose was administered by oral gavage.

Importantly, both lean and DIO mice treated with 6bK displayed significantly improved oral glucose tolerance compared to vehicle or inactive bisepi-6bK control groups (Fig. 3.3 and Fig. A3.9). Effects of similar magnitude on oral glucose tolerance in mice have been observed using several known human anti-diabetic therapeutics[33-36]. The two control groups exhibited similar blood glucose profiles, indicating that the observed effects of 6bK on glucose tolerance are lost when the stereochemistry of 6bK is altered in a way that abolishes IDE inhibition.

To further test if the observed effects of 6bK are specific to its interaction with IDE, we repeated these experiments using $IDE^{-/-}$ knockout mice[10,11]. Mice lacking IDE were unaffected by 6bK treatment, and exhibited blood glucose responses indistinguishable to that of the vehicle-treated cohort (Fig. A3.10). These results indicate that the effects of 6bK are dependent on the presence of IDE.

Collectively, these observations support a model in which IDE regulates glucose-induced insulin signaling, and therefore glucose tolerance, and demonstrate that acute IDE inhibition improves post-prandial glucose control in lean and DIO mice (Fig. 3.3a and 3.3b). Together, these results represent the first time that IDE inhibition has been shown to improve blood glucose tolerance[3,30].

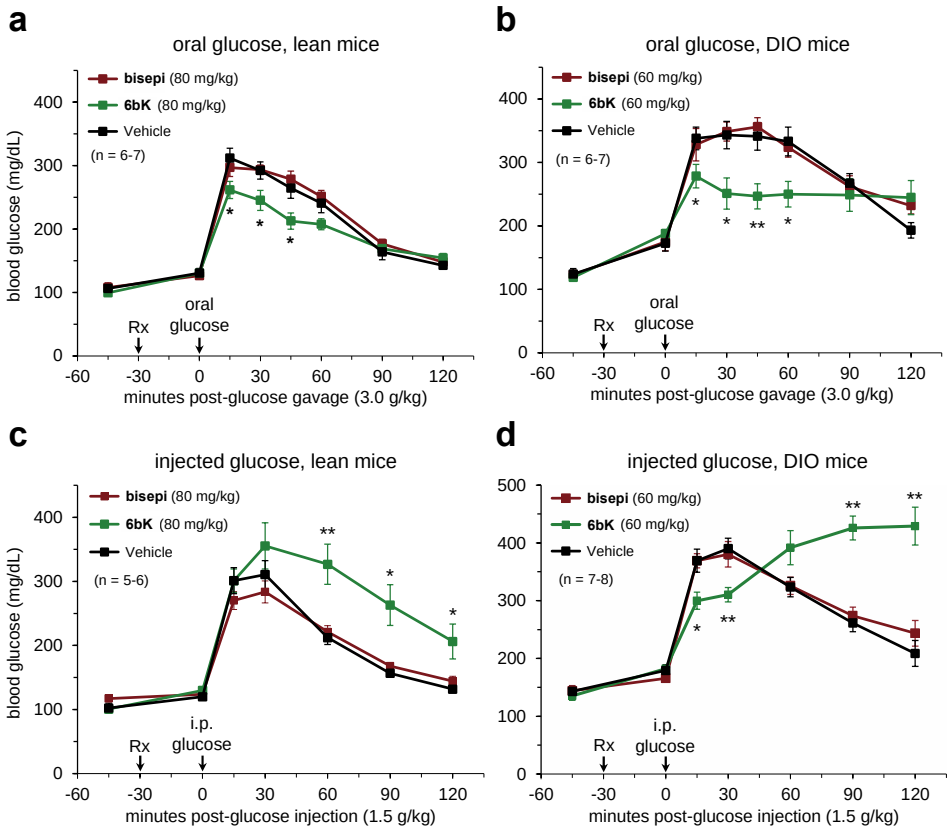


Figure 3.3 Physiological consequences of acute IDE inhibition by 6bK on glucose tolerance in lean and DIO mice. **a** and **b**, Oral glucose tolerance during acute IDE inhibition. **a**, Male C57BL/6J lean (25 g) mice were treated with a single i.p. injection of IDE inhibitor **6bK** (■, 80 mg/kg), inactive control **bisepi-6bK** (■, 80 mg/kg), or vehicle alone (■) 30 min prior to glucose gavage (3.0 g/kg). **b**, DIO mice (35-45 g) were treated with **6bK** (■, 60 mg/kg), and inactive control **bisepi-6bK** (■, 60 mg/kg) or vehicle alone (■) 30 min prior to glucose gavage (3.0 g/kg). **c** and **d**, Glucose tolerance phenotypes after i.p. injection of glucose (1.5 g/kg) in lean (**c**) and DIO (**d**) male mice treated with **6bK** (■), inactive **bisepi-6bK** (■), or vehicle alone (■). Area under the curve (AUC) calculations are shown in Extended Data Fig. 9. All data points and error bars represent mean ± SEM. Significance tests were performed using two-tail Student's t-test, and significance levels shown are $p < 0.05$ (*) or $p < 0.01$ (**) versus the vehicle-only control group.

3.6 IDE inhibition during an injected glucose challenge leads to impaired glucose tolerance

Prior work using $IDE^{-/-}$ mice characterized the effect of i.p. glucose injections, therefore we repeated the above experiments with 6bK followed by i.p. injected glucose tolerance tests (IPGTTs) to provide a more direct comparison with the knockout animal experiments[10,11]. In contrast to the observed improvement in oral glucose tolerance upon 6bK treatment (Figs. 3.3a and 3.3b), IDE inhibition with 6bK followed by a glucose injection (1.5 g/kg i.p.) resulted in *impaired* glucose tolerance after 2 h in both lean and obese mice compared to vehicle alone or bisepi-6bK-treated controls (Figs. 3.3c and 3.3d). These changes in the glucose response profiles of lean mice treated with 6bK compared to vehicle controls resemble the reported differences between $IDE^{-/-}$ and $IDE^{+/+}$ mice during similar IPGTTs[10,11]. Moreover, DIO mice treated with 6bK followed by glucose injection displayed a biphasic response in which glucose levels are lower over the initial 30 minutes of the IPGTT, followed by a hyperglycemic "rebound" starting 1 h after glucose injection (Fig. 3.3d). Both the suppression of peak glucose levels and the magnitude of the hyperglycemic rebound were dependent on 6bK dose, and neither effect was observed in cohorts treated with vehicle alone or inactive bisepi-6bK, indicating that the impaired glucose tolerance during an IPGTT correlates with IDE activity (Fig. 3.3, and Fig. A3.9).

Collectively, the results of the OGTTs and IPGTTs indicate that the route of glucose administration impacts the physiological response of 6bK-treated animals in ways that cannot be explained by a simple model in which IDE's physiological role is only to degrade insulin. Instead, the IPGTT results strongly suggest a role for IDE in regulating other glucose-regulating peptide hormones *in vivo*.

3.7 IDE regulates multiple hormones *in vivo*

The biochemical properties of IDE and its substrate recognition mechanism[23-25] enable this enzyme to cleave a wide range of peptide substrates *in vitro* (Appendix 3 Table 1). Two glucose-regulating hormones, beyond insulin, that are potential candidates for physiological regulation by IDE during a glucose challenge are glucagon and amylin. Purified IDE has been previously shown to cleave both peptides *in vitro*[37-39], but neither hormone is known to be regulated by IDE activity *in vivo*. Compared to insulin, glucagon is a modest *in vitro* IDE substrate (K_M = 3.5 μM for glucagon versus K_M < 30 nM for insulin)[37], although IDE is capable of degrading glucagon at a comparable rate if present in sufficiently high concentrations (k_{cat} = 38 min^{-1} for glucagon)[38]. Amylin is also a substrate for IDE *in vitro* (K_M = ~0.3 μM)[39]. Other proteases suggested to degrade glucagon include nardilysin, cathepsin B and D, in cells and *in vitro*[40,41], and neprilysin, which was shown to play a role in renal clearance of glucagon[42]. However, none of these enzymes are known to regulate endogenous processing of hormones or modulate blood glucose levels. To our knowledge, no proteases have been previously shown to degrade amylin *in vivo*[39].

To begin to probe the possibility that glucagon or amylin is regulated *in vivo* by IDE, we measured the plasma levels of these hormones at 20 and 130 minutes post-glucose injection in DIO mice treated with 6bK or vehicle alone during an IPGTT (Fig. 3.4a). Plasma collected 20 minutes post-glucose injection showed elevated insulin and amylin levels, but unchanged glucagon levels, for the 6bK-treated cohort relative to the control group (Fig. 2.4a). During the hyperglycemic rebound 130 min post-injection, glucagon levels for the 6bK group were strongly elevated above 200 pg/mL, compared with 90 pg/mL glucagon in control mice (Fig. 3.4a). Consistent with these elevated glucagon levels, expression of a gluconeogenesis transcriptional marker, G6Pase, was elevated in the livers of 6bK-treated mice compared to control mice (Fig. 3.4a).

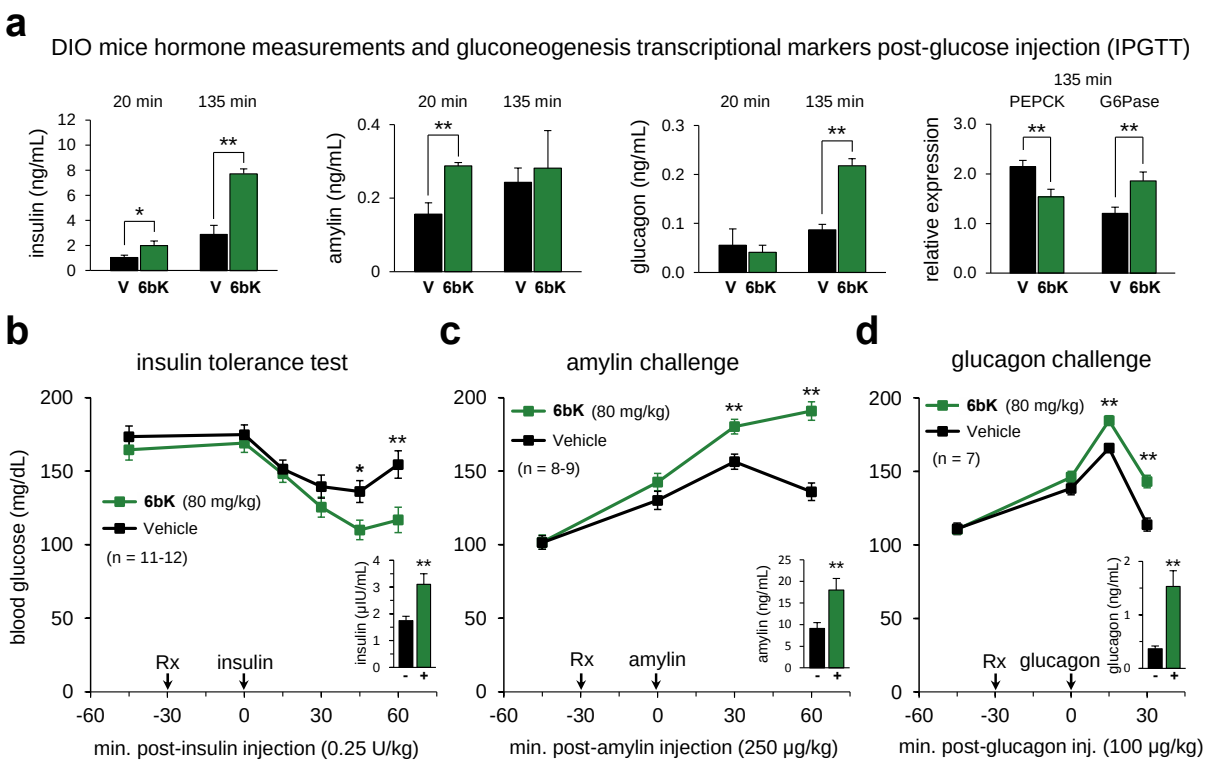


Figure 3.4 Acute IDE inhibition affects the abundance of multiple hormone substrates and their corresponding effects on blood glucose levels. **a**, Plasma hormone measurements at 20 and 135 minutes post-IPGTT (Fig. 3.3d) for DIO mice treated with 6bK (■) or vehicle alone (■). RT-PCR analysis of DIO liver samples collected at 135 min post-IPGTT reveals 50 % higher glucose-6-phosphatease (G6Pase) and 30 % lower phosphoenolpyruvate carboxykinase (PEPCK) transcript levels for the **6bK**-treated cohort (■) versus vehicle-only (**V**) controls (■). **b** to **d**, Blood glucose responses and abundance of injected hormones in lean mice 30 min after treatment with **6bK** (■, 80 mg/kg) or vehicle alone (■). **b**, Insulin s.c. (0.25 U/kg) after 5-hour fast. **c**, Amylin s.c. (250 µg/kg) after overnight fast. **d**, Glucagon s.c. (100 µg/kg) after overnight fast. Trunk blood was collected at the last time points for plasma hormone measurements (insets). All data points and error bars represent mean ± SEM. Significance tests were performed using two-tail Student's t-test, and significance levels shown are $p < 0.05$ (*) or $p < 0.01$ (**) versus the vehicle-only control group.

Because hormone abundance measurements can be difficult to interpret during fluctuations in blood glucose that in turn affect pancreatic hormone secretion, we performed additional studies to confirm the relationship between IDE activity and glucagon and amylin levels *in vivo*. To more directly establish the effect of IDE inhibition on the clearance of insulin, amylin, and glucagon *in vivo*, we injected each of these three hormones into lean mice 30 min after treatment with 6bK or vehicle alone (Figs. 3.4b to 3.4d). The 6bK-treated cohorts exhibited significantly stronger blood glucose responses to each of these hormones compared to vehicle controls; mice treated with 6bK experienced hypoglycemia during insulin tolerance tests (Fig. 3.4b) and hyperglycemia following challenges with either amylin (Fig. 3.4c) or glucagon (Fig. 3.4d) relative to control animals. Similar to glucagon, acute amylin administration in rodents results in a transient increase in blood glucose levels through gluconeogenesis and activation of lactic acid flux from muscle tissue to the liver[43]. Moreover, in each case the plasma level of the hormone injected remained elevated at the end of the procedure in 6bK-treated mice relative to control animals, demonstrating a role for IDE in regulating the abundance of these hormones (Figs. 3.4b to 3.4d insets).

3.8 IDE inhibition promotes glucagon signaling and gluconeogenesis

Taken together, these results strongly suggest that IDE activity regulates the stability and physiological activities of glucagon and amylin, in addition to insulin. Higher glucagon levels upon 6bK treatment provide a possible explanation for impaired glucose tolerance observed during an IPGTT. This model predicts that abrogating glucagon signaling should reverse the elevation of blood glucose by 6bK treatment during an IPGTT, while not substantially affecting the signaling pathways through which 6bK treatment lowers blood glucose during an OGTT.

To explicitly test these hypotheses, we repeated the glucose tolerance experiments using $GCGR^{-/-}$ mice that lack the G-protein coupled glucagon receptor (Fig. 3.5a and 3.5b)[44]. As

expected, mice lacking glucagon signaling exhibited an improvement in oral glucose tolerance upon 6bK treatment relative to vehicle controls that was similar to the oral glucose tolerance improvement observed in wild-type mice (Fig. 3.5a), consistent with a model in which insulin signaling in these mice is intact and regulated by IDE. In contrast, the responses of $GCGR^{-/-}$ mice treated with 6bK or vehicle alone to i.p. injected glucose were virtually identical, consistent with a model in which glucagon signaling drives the impaired glucose tolerance of wild-type lean and DIO mice upon 6bK treatment during an IPGTT (compare Figs. 3.5b and 3.3c). Collectively, these results reveal that IDE inhibition promotes glucagon signaling that can impair glucose tolerance during an IPGTT.

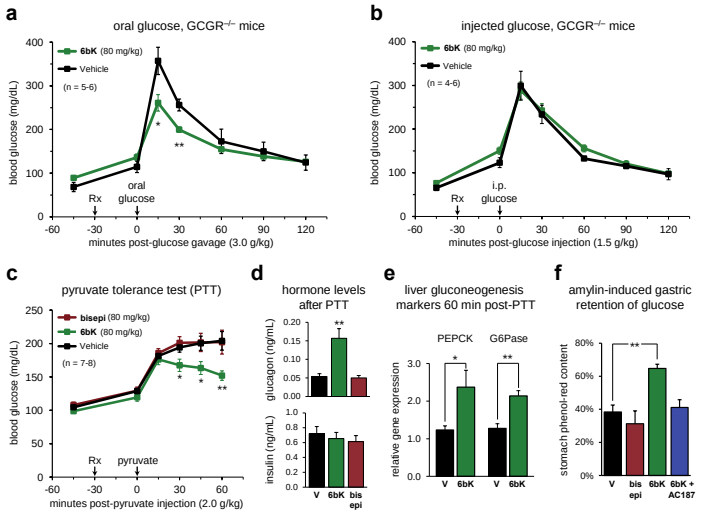


Figure 3.5 Acute IDE inhibition modulates the endogenous signaling activity of glucagon, amylin and insulin. **a** and **b**, G-protein-coupled glucagon receptor knockout mice (GCGR$^{-/-}$, C57BL/6J background) treated with IDE inhibitor **6bK** (■, 80 mg/kg) display altered glucose tolerance relative to vehicle-treated mice (■) if challenged with oral glucose (**a**) but not i.p. injected glucose (**b**). **c**, Wild-type mice fasted overnight were injected i.p. with pyruvate (2.0 g/kg) 30 min after treatment with **6bK** (■), inactive analog **bisepi-6bK** (■), or vehicle alone (■). **d**, Plasma hormone measurements 60 min post-PTT reveal elevated glucagon but similar insulin levels for the **6bK**-treated cohort (■) relative to **bisepi-6bK** (■), or vehicle (■) controls. **e**, RT-PCR analysis of liver samples 60-min post-PTT revealed elevated gluconeogenesis transcriptional markers for the **6bK**-treated group (■) relative to vehicle controls (■). **f**, Acute IDE inhibition slows gastric emptying through amylin signaling. Wild-type mice fasted overnight were given an oral glucose bolus (3.0 g/kg supplemented with 0.1 mg/mL phenol red) 30 min after treatment with **6bK** alone (■), **6bK** co-administered with the specific amylin receptor antagonist **AC187** (■, 3 mg/kg i.p.), vehicle alone (■), or inactive **bisepi-6bK** (■). The stomachs were dissected at 30 min post-glucose gavage. All data points and error bars represent mean ± SEM. Significance tests were performed using two-tail Student's t-test, and significance levels shown are $p < 0.05$ (*) or $p < 0.01$ (**) versus the vehicle-only control group.

To investigate the effect of IDE inhibition on endogenously secreted glucagon, we subjected mice treated with 6bK, vehicle, or inactive bisepi-6bK to a pyruvate tolerance test (PTT), which measures the ability of the liver under the action of glucagon to use pyruvate as a gluconeogenic substrate (Fig. 3.5c). The 6bK-treated cohort displayed significantly elevated plasma glucagon and increased expression of liver gluconeogenic markers compared to both control groups (Figs. 3.5d and 3.5e). The 6bK cohort also experienced significantly lower blood glucose during the PTT, suggesting that IDE inhibition produced a concomitant stimulation of glucose clearance that outweighed its effects on gluconeogenesis (Fig. 3.5c). These results collectively establish that IDE inhibition can augment endogenous glucagon signaling under conditions that favor gluconeogenesis.

3.9 IDE inhibition promotes amylin signaling and gastric emptying

Amylin is co-secreted with insulin, and is involved in glycemic control by inhibiting gastric emptying through the vagal route[45], promoting satiety during meals[46], and antagonizing glucagon signaling[47]. Pramlintide (Smylin) is a synthetic amylin derivative used clinically to control post-prandial glucose levels[33,36,48]. To determine the effects of IDE inhibition on endogenous amylin signaling, we measured gastric emptying efficiency[49], an amylin-specific process, in mice treated with 6bK, inactive control bisepi-6bK, or vehicle alone. Mice treated with 6bK exhibited 2-fold slower gastric emptying of a labeled glucose solution measured at 30 minutes post-gavage compared to vehicle and bisepi-6bK-treated controls (Fig. 3.5f). Importantly, co-administration of the specific amylin receptor antagonist AC187[50] blocked the effects of 6bK on gastric emptying (Fig. 3.5f). Collectively, these data reveal that IDE inhibition can slow post-prandial gastric emptying and demonstrate a role for IDE in modulating amylin signaling *in vivo*. These results also suggest that IDE inhibition may mimic the effect of amylin supplementation with pramlintide during meals[2,33].

3.10 Implications

The discovery and application of the first potent, highly selective, and physiologically active small-molecule IDE inhibitor revealed that acute IDE inhibition can lead to improved glucose tolerance in lean and obese mice after oral glucose administration, conditions that mimic the intake of a meal. These results validate the potential of IDE as a therapeutic target following decades of speculation[3,4]. Our data show that small-molecule IDE inhibitors can improve oral glucose tolerance to an extent comparable to that of DPP4 inhibitors[2,33]. The potential relevance of these animal studies to human disease is supported by the repeated recognition of IDE as a diabetes susceptibility gene in humans[5-11].

Equally important, additional *in vivo* and biochemical experiments using 6bK led to the discovery that IDE regulates the stability and signaling of glucagon and amylin, in addition to its established role in insulin degradation[10-12]. The identification of two additional pancreatic hormones as endogenous IDE substrates advances our understanding of the role of IDE in regulating physiological glucose homeostasis (Fig. 3.6). Amylin-mediated effects on gastric emptying and satiety during meals have been recently recognized to have therapeutic relevance in the treatment of diabetes[2,33], and our results represent the first demonstration of a small-molecule that can regulate both amylin and insulin signaling. Moreover, the link between IDE and glucagon revealed in this study provides additional evidence of the importance of glucagon regulation in human diabetes[51].

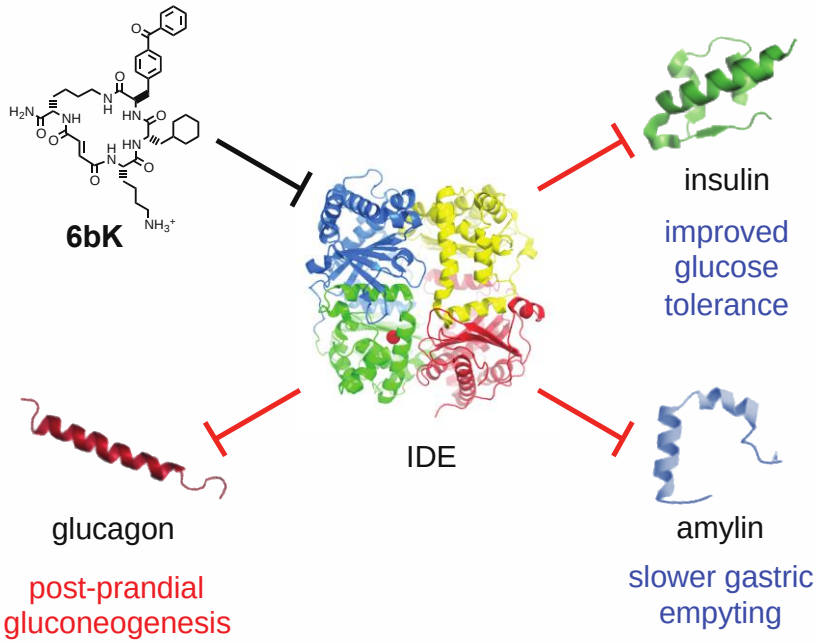


Figure 3.6 Model for the expanded roles of IDE in glucose homeostasis and gastric emptying based on the results of this study. IDE inhibition increases the abundance and signaling of three key pancreatic peptidic hormones, insulin, amylin, and glucagon, with the corresponding physiological effects shown in blue and red.

Our study also reveals a specific pharmacological requirement for therapeutic IDE inhibition—namely, that transient, rather than chronic, IDE inhibition is desirable to prevent elevation of glucagon signaling[51] (Fig. 3.6). A potential anti-diabetic therapeutic strategy motivated by these findings is the development of a fast-acting IDE inhibitor that can be taken with a meal to transiently augment endogenous insulin and amylin responses to help control post-prandial glycemia[33,34], and that is cleared or degraded before glucagon secretion resumes. Similar pre-meal therapeutic strategies with short-lived agents have already proven successful with fast-acting insulin analogs, secretagogues, and amylin supplementation[34,35]. It is tempting to speculate these agents may also have synergistic effects when co-administered with an IDE inhibitor[3,52]. Alternatively, the combination of an IDE inhibitor with incretin therapy[2,33] or a

glucagon receptor antagonist[51], may also prove therapeutic, as suggested by our experiments with glucagon-receptor deficient mice. Our findings also raise the possibility that future generations of IDE modulators may achieve substrate-selective inhibition[13] to enable the selective degradation of substrates such as glucagon while impairing the degradation of insulin.

The identification of a highly selective IDE inhibitor using a DNA-encoded library coupled with *in vitro* selection further establishes the value of this approach for the target-based discovery of bioactive small molecules. The unbiased nature of the selection led to the identification of a novel small-molecule binding site in IDE that enables unprecedented inhibition selectivity for this enzyme. These macrocycles and structural insights therefore may also prove useful in the development of assays that specifically target this binding site, and in the discovery of novel therapeutic leads that target IDE inhibition. More generally, this work highlights the value of physiologically active small-molecule probes to characterize the functions of genes implicated in human disease, an approach that could prove increasingly valuable as human genomic studies become more prevalent[15,16].

3.11 Methods

***In vitro* selection of a DNA-templated library with immobilized mouse IDE.** The *in vitro* selection used here was adapted from previously described protocols[1,2] using a DNA-templated library of 13,824 macrocycles[3]. Recombinant N-His_6-tagged mouse $IDE_{42\text{-}1019}$ (isoform containing the amino acids 42-1019 of the IDE sequence) was purified using immobilized cobalt magnetic beads (Dynabeads® His-Tag Isolation & Pulldown, Invitrogen®) according to the manufacturer's instructions. This purified IDE was confirmed to be catalytically active using the peptide substrate assay described below. IDE protein (~20 µg) was loaded onto the solid support by incubating the protein with beads (30 µL) at 25 °C for 30 min in 300 µL of pH 8.0 buffer containing 50 mM phosphate, 300 mM NaCl and 0.01 % Tween-20 (PBST buffer), and

washed twice with the same buffer. Two individually prepared protein-bead suspensions were incubated for 30 min with 5 pmol of the DNA-templated macrocycle library at RT, in pH 7.4 buffer containing 50 mM Tris-HCl, 150 mM NaCl, 0.05 % Tween-20 (TBST buffer) supplemented with 0.01 % BSA and 3 mg/mL yeast RNA (Ambion®). The beads were washed three times with 200 µL TBST buffer. The enriched library fraction was eluted by treatment with 200 mM imidazole in PBST buffer (50 µL) for 5 min.

The eluate solution was isolated and purified by buffer exchange using Sephadex spin-columns (Centrisep, Princeton Separation), according to the manufacturer's instructions. PCR amplification of the enriched pool of library barcodes was performed as previously reported[1], using primers that append adaptors for Illumina sequencing and a 7-base identifier. The long adaptor primer was 5'–AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGC TCTTCCGATCTXXXXXXCCCTGTACAC and the short adaptor primer was 5'–CAAGCAG AAGACGGCATACGAGCTCTTCCGATCTGAGTGGGATG (the 7-base identifier was XXXXXX). The PCR amplicons were purified by polyacrylamide gel electrophoresis, extracted, and quantified using qPCR and Picogreen assays (Invitrogen).

High-throughput DNA sequencing was performed on an Illumina Genome Analyzer instrument at the Harvard FAS Center for Systems Biology, Cambridge, MA, to yield an average of ~3.8 million sequence reads for each selection, untreated bead control and pre-selection library. Deconvolution of library barcodes and enrichment calculations were performed with custom software as described previously[1]. Variations in library member abundance as a result of binding to immobilized IDE was revealed by calculating fold-enrichment over the pre-selection library for the two independent selection experiments (Extended Data Fig. 1).

Protease assays with fluorogenic peptide substrates. The proteases $IDE_{42\text{-}1019}$, recombinant human $IDE_{42\text{-}1019}$ (R&D Systems), neprilysin (R&D), and angiotensin-converting enzyme (R&D) were assayed using the fluorophore/quencher-tagged peptide substrate Mca-RPPGFSAFK(Dnp)-OH (R&D) according to the manufacturer's instructions and recommended buffers. For IDE the recommended buffer is 50 mM Tris pH 7.5, 1 M NaCl. The enzyme mixtures (48 µL) were transferred to a 96-well plate and combined with 2 µL of inhibitor in DMSO solutions, in 3-fold dilution series. The mixtures were allowed to equilibrate for 10 min and the enzymatic reaction was started by addition of substrate peptide in assay buffer (50 µL), mixed, and monitored on a fluorescence plate reader (excitation at 320 nm, emission at 405 nm). Similarly, thimet oligopeptidase (R&D) and neurolysin (R&D) were assayed using substrate Mca-PLGPK(Dnp)-OH (R&D) according to the manufacturer's instructions and recommended buffers. Matrix metalloproteinase-1 (R&D) was activated and assayed according to the manufacturer's instructions with substrate Mca-KPLGL-Dpa-AR-NH_2 (R&D). All assay data points were obtained in duplicate. Data for the Yonetani–Theorell double inhibition plot was generated using 1 µL of each inhibitor (6b, and Ii1 or bacitracin) under otherwise identical conditions as above, at concentrations corresponding to ⅓x, 1x, 3x and 9x of respective IC_{50} values against hIDE (R&D).[4,5]

Protease assays with fluorogenic peptide substrates.IDE degradation assays for insulin and calcitonin-gene related peptide (CGRP). A solution of 0.4 µg/mL IDE (R&D) in pH 7.5 buffer containing 50 mM Tris, 1.0 M NaCl (48 µL) was transferred to a 96-well plate, and combined with 2 µL of each inhibitor (6b, 6bK and 28) dissolved in DMSO, in 3-fold dilution series (Extended Data Fig. 2c). A solution of insulin (50 µL) was added to a final concentration of 10 ng/mL, and incubated at 30 °C for 15 min. This procedure was optimized to result in ~75 % degradation of insulin. The reaction was terminated by addition of inhibitor 6bK to a final

concentration of 20 µM and chilled on ice. Insulin was quantified using 10 µL of the enzymatic reaction for the sensitive-range protocol Homogeneous Time-Resolved FRET Insulin assay (CysBio®) in 20 µL total volume according to the manufacturer's instructions. Fluorescence was measured using a Tecan Safire 2 plate reader (excitation = 320 nm, emission = 665 and 620 nm, lag time = 60 µs) according the assay manufacturer's recommendations. Degradation of CGRP by endogenous IDE in mouse plasma was analyzed by LC-MS for the formation of $CGRP_{1\text{-}17}$ and $CGRP_{18\text{-}37}$ as previously reported[6]. The substrate was added to a final concentration of 10 µM, and 6b was added to a final concentration of 0.1 to 10 µM (Extended Data Fig. 2d).

General procedure for synthesis of macrocycle inhibitors. Rink amide resin (NovaPEG Novabiochem®, ~0.49 mmol/g, typically at a scale of 0.1 to 2 mmol) was swollen with ~10 volumes of anhydrous DMF for 1 h in a peptide synthesis vessel with mixing provided by dry nitrogen bubbling. In a separate flask, N^{α}-allyloxycarbonyl-N^{ε}-2-Fmoc-L-lysine (5 equiv.) and 2-(1*H*-7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyl uronium hexafluorophosphate (HATU, 4.75 equiv.) were dissolved in anhydrous DMF (~10 vol.), then treated with *N,N'*-diisopropylethylamine (DIPEA, 10 equiv.) for 5 min at RT. The solution was combined with the pre-swollen Rink amide resin and mixed with nitrogen bubbling overnight. The vessel was eluted and the resin was washed three times with *N*-methyl-2-pyrrolidone (NMP, ~10 vol.). Following each coupling step, Fmoc deprotection was effected with 20 % piperidine in NMP (~20 vol.) for 20 min, repeated three times, followed by washing three times with NMP (~10 vol.) and twice with anhydrous DMF (~10 vol.). The general procedure for amide coupling of building blocks A, B and C was treatment of the resin with solutions of HATU-activated N^{α}-Fmoc amino acids (5 equiv.) for 3-5 hours in anhydrous DMF, mixing with dry nitrogen bubbling. The general

procedure for HATU-activation was treating a solution of N^{α}-Fmoc amino acid (5 equiv.) and HATU (4.75 equiv.) in anhydrous DMF (10 vol.) with DIPEA (10 equiv.) for 5 min at RT.

Following the final Fmoc deprotection procedure, the α-amine of building block C was coupled with allyl fumarate monoester (10 equiv.) using activation conditions as previously described with HATU (9.5 equiv.) and DIPEA (20 equiv.) in anhydrous DMF (~10 vol.). Allyl fumarate coupling was accomplished by overnight mixing with dry nitrogen bubbling, followed by washing five times with NMP (~10 vol.) and three times with $CHCl_3$ (~10 vol.). Simultaneous allyl ester and *N*-allyloxycarbonyl group cleavage in solid support was effected with three consecutive treatments with a solution of tetrakis(triphenylphosphine)palladium(0) (0.5 equiv.) dissolved in degassed $CHCl_3$ containing acetic acid and *N*-methylmorpholine (40:2:1 ratio, ~20 vol.), mixing by nitrogen bubbling for 30 min. The resin was then washed twice subsequently with ~20 vol. of 5 % DIPEA in DMF, then twice with a 5 % solution of sodium diethyldithiocarbamate trihydrate in DMF (~20 vol.), twice with 5 % solution of hydroxybenzotriazole monohydrate in DMF, and finally washed with 50 % CH_2Cl_2 in DMF and re-equilibrated with anhydrous DMF (~10 vol.).

Treating the resin with pentafluorophenyl diphenylphosphinate (5 equiv.) and DIPEA (10 equiv.) in anhydrous DMF (~10 vol.) mixing by nitrogen bubbling overnight produced the macrocyclized products. The resin was washed with NMP (~20 vol.) and CH_2Cl_2 (~20 vol.) and dried by vacuum. The macrocyclized product was cleaved from the resin by two 15 min treatments of the macrocycle-bound resin with 95 % TFA containing 2.5 % water and 2.5 % triisopropylsilane (~20 vol.), followed by two TFA washes (~5 vol.). The TFA solution was dried to a residue under rotatory evaporation, and the peptide was precipitated by the addition of dry Et_2O. The ether was decanted and the remaining solid was dried and dissolved in a minimum volume of 3:1 DMF-water prior to purification by liquid chromatography. HPLC purification was performed on a C18 21.2x250 mm column (5 µm particle, 100 Å pore size, Kromasil®), using a gradient of 30 % to 80 % MeCN/water over 30 min, and solvents containing 0.1 % TFA.

Fractions containing the desired macrocyclic peptide were combined and freeze-dried to produce a white powder. Typical yields were 5-15 % based on resin loading. Purity was determined by HPLC (Zorbax SB-C18 2.1x150 mm column, 5 µm particle) with UV detection at 230 nm, using a gradient of 30 % to 80 % MeCN/water over 30 min, and solvents containing 0.1 % TFA. The formula of final products was confirmed by accurate mass measurements using an Agilent 1100 LC-MSD SL instrument (Supplementary Table S2).

Dosing and formulation of macrocycle inhibitors for *in vivo* studies. The doses of 6bK used in this study were chosen based on literature precedent for small molecules and drugs of similar potencies[7-13]. Procedures using experimental compounds were in accordance with the standing committee on the Use of Animals in Research and Teaching at Harvard University, the guidelines and rules established by the Faculty of Arts and Sciences' Institutional Animal Care and Use Committee (IACUC), and the National Institutes of Health Guidelines for the Humane Treatment of Laboratory Animals.

Purified macrocycle inhibitors were dissolved in DMSO-d_6 (~200 to 250 mg/mL stock solutions). A sample aliquot (5 µL) of the macrocycle solution was diluted with 445 µL DMSO-d_6, then combined with 50 µL of freshly prepared solution of 20 mM CH_2Cl_2 in DMSO-d_6 in for ^{1}H-NMR acquisition (600 MHz, relaxation time = 2 sec). The inhibitor concentration was calculated using the integral of the CH_2Cl_2 singlet (δ 5.76 ppm, 2H)[14], which appears in an uncluttered region of these spectra[1,15]. For injectable formulations (10 mL/kg i.p. injection volume), the macrocycle inhibitor solution in DMSO-d_6 (200 – 250 mg/mL, based on free-base molecular weight) was combined with 1:20 w/w Captisol® powder (CyDex)[16]. The resulting slurry was supplemented with DMSO-d_6 to make up to 5 % of the final formulation volume, mixed thoroughly and dissolved with sterile saline solution (0.9 % NaCl). Vehicle controls were identically formulated with 5 % DMSO-d_6 and equal amount of Captisol®. The formulated

solutions of inhibitor were clear with no visible particles, and were stored overnight at 4 °C prior to injection.

Expression and purification of recombinant cysteine-free hIDE (IDE-CF). We expressed cysteine-free catalytically-inactive, human IDE (IDE-CF) in pPROEX vector[17]. IDE-CF contains the following substitutions: C110L, E111Q, C171S, C178A, C257V, C414L, C573N, C590S, C789S, C812A, C819A, C904S, C966N, and C974A. IDE was expressed and purified as previously described[17]. Briefly, IDE-CF was transformed into *E. coli* BL21-CodonPlus (DE3)-RIL, grown at 37 °C to a cell density of 0.6 O.D. and induced with IPTG at 16 °C for 19 hours. Cells were lysed and the lysate was subjected to Ni-affinity (GE LifeScience) and anion exchange chromatography (GE LifeScience). The protein was further purified by size exclusion chromatography (Superdex S200 column) three successive times, first without inhibitor then two times after addition of 2-fold molar excess of 6b.

IDE-CF•6b co-crystallization. Eluent from size exclusion chromatography was concentrated to 20 mg/ml in 20 mM Tris, pH 8.0, 50 mM NaCl, 0.1 mM PMSF and 2-fold molar excess of 6b was added to form the protein-inhibitor complex. The complex was mixed with equal volumes of reservoir solution containing 0.1 M HEPES (pH 7.5), 20 % (w/v) PEGMME-5000, 12 % tacsimate and 10 % dioxane. Crystals appeared after 3-5 days at 25 °C and were then equilibrated in cryoprotective buffer containing well solution and 30 % glycerol. IDE-CF•6b complex crystals belong to the space group $P6_5$, with unit-cell dimensions $a = b = 262$ Å and $c = 90$ Å, and contain two molecules of IDE per asymmetric unit (Extended Data Figure 3a).

X-Ray Diffraction. X-ray diffraction data were collected at the National Synchrotron Light Source at Brookhaven National Laboratories beamline X29 at 100K and 1.075 Å.

IDE-CF•6b structure determination. Data were processed in space group $P6_5$ using autoProc[18]. The structure was phased by molecular replacement using the structure for human IDE E111Q (residues 45-1011, with residues 965-977 missing) in complex with inhibitor compound 41367 (PDB: 2YPU) as a search model in Phaser[19]. The model of the structure was built in Coot[20] and refined in PHENIX[21], using NCS (torsion-angle) and TLS (9 groups per chain). In the Ramachandran plot, 100 % of the residues appear in the allowed regions, 97.2 % of the residues appear in the favored regions, and 0 % of the residues appear in the outlier regions (Extended Data Figure 3a). The structural model of IDE and inhibitor was refined using non-crystallographic symmetry restraints due to the high degree of structural similarity between the two chains of the protein the asymmetric unit. The root-mean-square deviation between the two chains is 0.5 Å for the protein chains, and 0.9 Å for the ligand atoms, indicating high structural similarity between the two chains. Structure coordinates are deposited in the Protein Data Bank (accession number 4TLE).

Macrocycle docking simulations. Receptor and ligand preparation was performed in the standard method[22,23]. DOCKing was performed using version 6.6 with default parameters for flexible ligand and grid-based scoring, and the van der Waals exponent was 9. Because of the mutagenesis data strongly pointing to a role of Ala479, we limited docking of the macrocycle to an area within 15 Å of Ala479. The highest scoring poses (Extended Data Fig. 3), by both gridscore5 and Hawkins GB/SA6, are consistent with the placement of building blocks A (benzophenone), B (cyclohexyl ring), and *trans* olefin of 6b in the proposed structure. Docking calculations for the inactive isomer 6a were unsuccessful, while for the low-activity analog 11

the result was a different and lower-scoring pose than 6b, consistent with the mutagenesis and biochemical assay results (Table S5, Fig. 1b, Extended Data Fig. 4).

Anisotropy binding assay. Catalytically-inactive CF-IDE was titrated with fluorescein-labeled macrocycle 31 in 20 mM Tris buffer pH 8.0, with 50 mM NaCl, and 0.1 mM PMSF at 25°C in the presence or absence of insulin (2.15 µM). The final assay volume was 220 µL, with a final DMSO concentration of 0.1 % and a final 31 concentration of 0.9 nM. After equilibration, the increase in the fluorescence anisotropy of the fluorescent ligand was recorded at 492 nm, using excitation at 523 nm, and fitted against a quadratic binding equation in Kaleidagraph (Synergy Software) to yield the dissociation constant (K_D = 43 nM). From our structural work, we hypothesized that insulin and macrocycle 31 compete for the same binding site.

Equation 1[24] approximates how the concentration of insulin ([I]) and the affinity constant of insulin (K_i = 280 nM) affect the apparent dissociation constant (K_D^{app}) of macrocycle 31 for CF-IDE. In the absence of competing insulin, 31 binds with the dissociation constant (K_D = 43 nM) to CF-IDE. In the presence of 2.15 uM insulin, the calculated K_D^{app} for macrocycle 31 is 363 nM (~8-fold shift) which is in good agreement with the competitive binding mode revealed by the crystal structure, and the anisotropy measurements (observed ~7-fold shift, K_D^{app}= ~ 280 nM, Extended Data Fig. 3e).

$$K_{\mathrm{D}}^{\mathrm{app}} = \left(1 + \frac{[\mathrm{I}]}{K_{\mathrm{i}}} \right) . K_{\mathrm{D}}$$

Eq. 1

Site-directed mutagenesis, expression, and purification of human IDE. The reported N-His_6-tagged human $IDE_{42\text{-}1019}$ construct was introduced in the expression plasmid pTrcHis-A (Invitrogen) using primers for uracil-specific excision reactions (USER) by *Taq* (NEB) and *Pfu* polymerases (PfuTurbo CX®, Agilent). The IDE gene was amplified with the primers 5'-

ATCATCATATGAATAATCCAGCCA-*dU*-CAAGAGAATAGG and 5'-ATGCTAGCCATACC*TCAGA G-dU-TTTGCAGCCATGAAG* (underlined sequences represent overhangs, and italics highlight the PCR priming sequence). Similarly, the pTrcHis-A vector was amplified for USER cloning with the primers 5'-ATGGCTGGATTATTCAT*ATGATGA-dU-GATGATGATGAGAACCC* and 5'-ACTCTGA*GGTATGGCTAGCA-dU-GACTGGTG*. Mutant IDE constructs were generated by amplifying the full vector construct with USER cloning primers introducing a mutant overhang (Supplementary Table S3).

All PCR products were purified on microcentrifuge membrane columns (MinElute®, Quiagen) and quantified by UV absorbance (NanoDrop). Each fragment (0.2 pmol) was combined in a 10 µL reaction mixture containing 20 units *Dpn*I (NEB), 0.75 units of USER mix (Endonuclease VIII and Uracil-DNA Glycosylase, NEB), 20 mM Tris-acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 1 mM dithiothreitol at pH 7.9 (1x NEBuffer 4). The reactions were incubated at 37 °C for 45 min, followed by heating to 80 °C and slow cooling to 30 °C (0.2 °C/s). The hybridized constructs were directly used for heat-shock transformation of chemically competent NEB turbo *E. coli* cells according to the manufacturer's instructions. Transformants were selected on carbenicillin LB agar, and isolated colonies were cultured overnight in 2 mL LB.

The plasmid was extracted using a microcentrifuge membrane column kit (Miniprep®, Qiagen), and the sequence of genes and vector junctions were confirmed by Sanger sequencing (Supplementary Table S4). The plasmid constructs were transformed by heat-shock into chemically-competent expression strain Rosetta 2 (DE3) pLysS *E. coli* cells (EMD Millipore), and selected on carbenicillin/chloramphenicol LB agar. Cells transformed with IDE pTrcHis A constructs were cultured overnight at 37 °C in 2 XYT media (31 g in 1L) containing 100 µg/mL ampicillin and 34 µg/mL chloramphenicol. Expression of His6-tagged IDE proteins was induced when the culture measured OD600 ~0.6 by addition of isopropyl-β-D-1-

thiogalactopyranoside (IPTG) to 1 mM final concentration, incubated overnight at 37 °C, followed by centrifugation at 10,000 *g* for 30 min at 4 °C.

Recombinant His6-tagged proteins were purified by Ni(II)-affinity chromatography (IMAC sepharose beads, GE Healthcare®) according to the manufacturer's instructions. The cell pellets were resuspended in pH 8.0 buffer containing 50 mM phosphate, 300 mM NaCl, 10 mM imidazole, 1 % Triton X-100 and 1 mM tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and were lysed by probe sonication for 4 min at <4 °C, followed by clearing of cell debris by centrifugation at 10,000 *g* for 25 min at 4 °C. The supernatant was incubated with Ni(II)-doped IMAC resin (2 mL) for 3 h at 4 °C. The resin was washed twice with the cell resuspension/lysis buffer, and three times with pH 8.0 buffer containing 50 mM phosphate, 300 mM NaCl, 50 mM imidazole and 1 mM TCEP. Elution was performed in 2 mL aliquots by raising the imidazole concentration to 250 mM and subsequently to 500 mM in the previous buffer. The fractions were combined and the buffer was exchanged to the recommended IDE buffer (R&D) using spin columns with 100 KDa molecular weight cut off membranes (Millipore). Protein yields were typically ~10 µg/L, and >90 % purity based on gel electrophoresis analysis (Coomassie stained). IDE-specific protease activity was >95 % as assessed by inhibition of degradation of peptide substrate Mca-RPPGFSAFK(Dnp)-OH (R&D) by 20 µM 6bK, compared with pre-quantitated commercially available human IDE (R&D). The complete list of IDE mutations assayed is shown in Supplementary Table S5.

***In vivo* studies, general information.** Wild-type C57BL/6J and diet-induced obese (DIO) C57BL/6J age-matched male adult mice were purchased from Jackson Laboratories. The age range was 13 to 15 weeks for lean mice, and 24 to 26 weeks for DIO mice. All animals were individually housed on a 14-h light, 10-h dark schedule at the Biology Research Infrastructure (BRI), Harvard University. Cage enrichment included cotton bedding and a red plastic hut. Water and food were available *ad libitum*, consisting respectively of normal chow

(Prolab® RHM 3000) or high-fat diet (60 kcal % fat, D12492, Research Diets Inc.). Adult IDE knock-out mice ($IDE^{-/-}$) and control mice ($IDE^{+/+}$) were obtained from Mayo Clinic (Florida), and housed in the Biology Research Infrastructure (BRI), Harvard University, for 8 weeks prior to experiments as described above. Experiments were conducted on $IDE^{-/-}$ and $^{+/+}$ age-matched mice cohorts ranging 17 to 21 weeks old. All animal care and experimental procedures were in accordance with the standing committee on the Use of Animals in Research and Teaching at Harvard University, the guidelines and rules established by the Faculty of Arts and Sciences' Institutional Animal Care and Use Committee (IACUC), and the National Institutes of Health Guidelines for the Humane Treatment of Laboratory Animals. Glucagon-receptor knock-out mice ($GCGR^{-/-}$) and control WT mice ($GCGR^{+/+}$) were group-housed with a 14-h light and 10-h dark schedule and treated in accordance with the guidelines and rules approved by the IACUC at Albert Einstein College of Medicine, NY. Power analysis to determine animal cohort numbers was based on preliminary results and literature precedent, usually requiring between 5 and 8 animals per group. Animals were only excluded from the cohorts in cases of chronic weakness, which occurs among $GCGR^{-/-}$ mice, or when we identified occasional DIO mice with an outlier diabetic phenotype (e.g. >200 mg/dL fasting blood glucose). Age- and weight-matched mice were randomized to each treatment group. Double-blinding was not feasible.

Glucose tolerance tests GTT and blood glucose measurements. Prior to a glucose challenge, the animals were fasted for 14 h (8 pm to 10 am, during the dark cycle) while individually housed in a clean cage with inedible wood-chip as a floor substrate, cotton bedding and a red plastic hut. Inhibitor, vehicle or control compounds were administered by a single intraperitoneal (i.p.) injection 30 min prior to the glucose challenge. Dextrose was formulated in sterile saline (3 g in 10 mL total), and the dose was adjusted by fasted body weight. For oGTT, 3.0 g/kg dextrose was administered by gavage at a dose of 10 mL/kg, and for ipGTT, 1.5 g/kg

dextrose injected at a dose of 5 mL/kg. Blood glucose was measured using AccuCheck® meters (Aviva) from blood droplets obtained from a small nick at the tip of the tail, at timepoints -45, 0, 15, 30, 45, 60, 90 and 120 min with reference to the time of glucose injection. The area of the blood glucose response profile curve corresponding to each animal was calculated by the trapezoid method[25], using as reference each individual baseline blood glucose measurement prior to glucose administration (t = 0). The sum of the trapezoidal areas between the 0, 15, 30, 45, 60, 90 and 120 minute time points corresponding to each animal were summed to obtain the area under the curve (AUC). The relative area values are expressed as a percentage relative to the average AUC of the vehicle cohort, which is defined as 100 % (Extended Data Fig 9). Values are reported as mean ± S.E.M. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures are * $p < 0.05$ *versus* vehicle control group or ** $p < 0.01$ *versus* vehicle control group.

Insulin Tolerance Test (ITT), Glucagon Challenge (GC) and Amylin Challenge (AC)

For hormone challenges animals were fasted individually housed as described above. For ITT the fasting period was 6 h (7 am to 1 pm), and for glucagon and amylin challenges the fasting period was 14 h (8 pm – 10 am). Inhibitor or vehicle alone was injected i.p. as previously described, 30 min prior to each hormone challenge. Insulin (Humulin-R®, Eli Lilly) was injected subcutaneously (s.c.) 0.25 U/kg formulated in sterile saline (5 mL/kg). Glucagon (Eli Lilly) was injected s.c. 100 µg/kg formulated in 0.5 % BSA sterile saline (5 mL/kg). Amylin (Bachem) was injected s.c. 250 µg/kg formulated in sterile saline (5 mL/kg). Blood glucose was measured at timepoints -45, 0, 15, 30, 45, 60 and 75 min with reference to the time of hormone injection, by microsampling from a tail nick as described above. Values are reported as mean ± S.E.M. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures are * $p < 0.05$ *versus* vehicle control group or ** $p < 0.01$ versus vehicle control group.

Blood collection and plasma hormone measurements. Blood was collected in EDTA-coated tubes (BD Microtainer®) from trunk bleeding (~500 µL) after CO_2-euthanasia for all hormone assays. The plasma was immediately separated from red blood cells by centrifugation 10 min at 1800 *g*, aliquoted, frozen over dry ice and stored at -80 °C. Insulin, glucagon, amylin and pro-insulin C-peptide fragment were quantified from 10 µL plasma samples using magnetic-bead Multiplexed Mouse Metabolic Hormone panel (Milliplex, EMD Millipore) according to the manufacturer's instructions, using a Luminex FlexMap 3D instrument. Plasma containing high levels of human insulin (Humulin-R) were quantified using 25 µL samples with Insulin Ultrasensitive ELISA (ALPCO). Values are reported as mean ± S.E.M. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures are * $p < 0.05$ *versus* vehicle control group or ** $p < 0.01$ versus vehicle control group.

Gastric emptying measurements. Mice were fasted 14 h (8 pm to 10 am). Inhibitor or vehicle alone was injected i.p. as previously described, followed 30 min later by an oral glucose bolus (3.0 g/kg, 10 mL/kg) in sterile saline containing 0.1 mg/mL phenol red. At 30 min, the stomach was promptly dissected after CO_2-euthanasia and stored on ice. The stomach contents were extracted into 2.5 mL EtOH (95 %) by homogenization for 1 min using a probe sonicator. Tissue debris were decanted by centrifugation at 4000 *g*, followed by clearing at 15000 *g* for 15 min. The supernatant (1 mL) was mixed with 0.5 mL of aqueous NaOH (20 mM), and incubated at -20 °C for 1 h. The solution was centrifuged at 15,000 *g* for 15 min, and absorbance was determined at 565 nM. The spectrophotometer was blanked with the stomach contents of a mouse treated with colorless glucose solution. The absorbance was adjusted to the amount of glucose solution dosed to each mouse. Values are reported as mean ± S.E.M relative to the original phenol red glucose solution. Statistics were performed using a two-tail

Student's t-test, and significance levels shown in the figures are * $p < 0.05$ *versus* vehicle control group or ** $p < 0.01$ *versus* vehicle control group.

Stable isotope dilution LC-MS, pharmacokinetics, and tissue distribution measurements. Heavy-labeled macrocycle inhibitor (heavy-6bK, Extended Data Fig. 5) was synthesized as described above, substituting "building block C" with N^{α}-Fmoc-N^{ε}-Boc-lysine $^{13}C_6$ $^{15}N_2$ (98 atom %, Sigma-Aldrich). The product was 8 mass-units heavier than 6bK, otherwise with identical properties and IC_{50}. Plasma samples (15 µL) from 6bK-treated mice and vehicle controls were combined with 5 µL of heavy-6bK in PBS (final concentration of 10 µM), and incubated for 30 min on ice. Plasma proteins were precipitated with 180 µL cold 1 % TFA in MeCN, sonicated 2 min, and centrifuged 13,000 *g* for 1 min. The supernatant was diluted 100- and 1000-fold for liquid chromatography-mass spectrometry (LC-MS) analysis. Tissue samples (~100 mg) from 6bK-treated mice and vehicle controls were weighed and disrupted in Dounce homogenizers with PBS buffer (0.5 mL/100 mg sample), supplemented with 5 µM heavy-6bK and protease inhibitor cocktail (1 tablet/50 mL PBS, Roche diagnostics). The lysate was incubated on ice for 30 min, sonicated 5 min and centrifuged at 13,000 *g* for 5 min. The bulk of supernatant proteins were precipitated by denaturation at 95 °C for 5 min, removed by centrifugation. A 50 µL aliquot of the supernatant was treated with 450 µL cold 1 % TFA in MeCN, cleared by centrifugation and diluted 100-fold for LC-MS analysis as described above for plasma samples. A standard curve of 6bK and heavy-6bK (1 µM each and 3-fold serial dilutions) were used for LC-MS quantitation using a Waters Q-TOF premier instrument.

RT-PCR analysis of liver gluconeogenesis markers phosphoenolpyruvate carboxykinase (PEPCK) and glucose-6-phosphatase (G6Pase). Total RNA was isolated from liver samples (~100 mg) using TRIzol® reagent (Invitrogen, 1 mL/100 mg sample), followed by spin-column purification (RNeasy® kit, Qiagen) and on-column DNAse treatment

(Qiagen) according to the manufacturer's instructions. RNA concentrations were determined by UV spectrophotometry (NanoDrop). One microgram of total RNA was used for reverse transcription with oligo(dT) primers (SuperScript® III First-Strand Synthesis SuperMix, Life Technologies) according to the manufacturer's instructions. Quantitative PCR reactions included 1 µL of cDNA diluted 1:100, 0.4 µM primers, 2x SYBR Green PCR Master Mix (Invitrogen) in 25 µL total volume, and were read out by a CFX96™ Real-Time PCR Detection System (BioRad). Transcript levels were determined using two known primer pairs[26,27] for each gene of interest (Supplementary Table S4), which were normalized against tubulin and beta-actin transcripts ($\Delta\Delta C_T$ method), and expressed relative to the lowest sample. Duplicate control assays without reverse transcriptase treatment were run for each RNA preparation and primer set used. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures are * $p < 0.05$ *versus* vehicle control group or ** $p < 0.01$ *versus* vehicle control group.

Amyloid peptide measurements. Lean C57BL/6J mice were treated with 6bK (80 mg/kg), vehicle alone, or bisepi-6bK (80 mg/kg). Two hours post-injection, hemibrains were promptly dissected after CO_2-euthanasia and stored at -80 °C until extraction. A hemibrain (~150 mg wet weight) was homogenized in a Dounce homogenizer in 0.2% diethylamine and 50 mM NaCl (900 µL). The homogenate was centrifuged at 20,000 × g for 1 h at 4°C to remove insoluble material. A fraction of supernatant (100 µL) was neutralized with 1:10 volume of Tris HCl, pH 6.8. The sample was diluted 1:4 in assay buffer, and analyzed for Aβ(40) and Aβ(42) levels using the respective Aβ ELISA assays (Invitrogen) following the manufacturer's protocols.

3.12 References

1 Drag, M. & Salvesen, G. S. Emerging principles in protease-based drug discovery. *Nat Rev Drug Discov* 9, 690-701, doi:nrd3053 [pii]10.1038/nrd3053 (2010).

2 Drucker, D. J. The biology of incretin hormones. *Cell Metab* 3, 153-165, doi:S1550-4131(06)00028-3 [pii]10.1016/j.cmet.2006.01.004 (2006).

3 Duckworth, W. C., Bennett, R. G. & Hamel, F. G. Insulin degradation: progress and potential. *Endocr Rev* 19, 608-624 (1998).

4 Mirsky, I. A. & Broh-Kahn, R. H. The inactivation of insulin by tissue extracts; the distribution and properties of insulin inactivating extracts. *Arch Biochem* 20, 1-9 (1949).

5 Karamohamed, S. *et al.* Polymorphisms in the insulin-degrading enzyme gene are associated with type 2 diabetes in men from the NHLBI Framingham Heart Study. *Diabetes* 52, 1562-1567 (2003).

6 Gu, H. F. *et al.* Quantitative trait loci near the insulin-degrading enzyme (IDE) gene contribute to variation in plasma insulin levels. *Diabetes* 53, 2137-2142 (2004).

7 Sladek, R. *et al.* A genome-wide association study identifies novel risk loci for type 2 diabetes. *Nature* 445, 881-885, doi:10.1038/nature05616 (2007).

8 Zeggini, E. *et al.* Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science* 316, 1336-1341, doi:10.1126/science.1142364 (2007).

9 Bartl, J. *et al.* Disorder-specific effects of polymorphisms at opposing ends of the Insulin Degrading Enzyme gene. *BMC medical genetics* 12, 151, doi:10.1186/1471-2350-12-151 (2011).

10 Farris, W. *et al.* Insulin-degrading enzyme regulates the levels of insulin, amyloid beta-protein, and the beta-amyloid precursor protein intracellular domain in vivo. *Proc Natl Acad Sci U S A* 100, 4162-4167, doi:10.1073/pnas.0230450100 [pii] (2003).

11 Abdul-Hay, S. O. *et al.* Deletion of insulin-degrading enzyme elicits antipodal, age-dependent effects on glucose and insulin tolerance. *PLoS One* 6, e20818, doi:10.1371/journal.pone.0020818 (2011).

12 Leissring, M. A. *et al.* Designed inhibitors of insulin-degrading enzyme regulate the catabolism and activity of insulin. *PLoS One* 5, e10504, doi:10.1371/journal.pone.0010504 (2010).

13 Abdul-Hay, S. O. *et al.* Optimization of peptide hydroxamate inhibitors of insulin-degrading enzyme reveals marked substrate-selectivity. *J Med Chem* 56, 2246-2255, doi:10.1021/jm301280p (2013).

14 Saghatelian, A., Jessani, N., Joseph, A., Humphrey, M. & Cravatt, B. F. Activity-based probes for the proteomic profiling of metalloproteases. *Proc Natl Acad Sci U S A* 101, 10000-10005, doi:10.1073/pnas.0402784101 (2004).

15 Knight, Z. A. & Shokat, K. M. Chemical genetics: where genetics and pharmacology meet. *Cell* 128, 425-430, doi:10.1016/j.cell.2007.01.021 (2007).

16 Workman, P. & Collins, I. Probing the probes: fitness factors for small molecule tools. *Chem Biol* 17, 561-577, doi:10.1016/j.chembiol.2010.05.013 (2010).

17 Gartner, Z. J. *et al.* DNA-templated organic synthesis and selection of a library of macrocycles. *Science* 305, 1601-1605, doi:10.1126/science.1102629 [pii] (2004).

18 Tse, B. N., Snyder, T. M., Shen, Y. & Liu, D. R. Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *J Am Chem Soc* 130, 15611-15626, doi:10.1021/ja805649f (2008).

19 Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K. & Liu, D. R. In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J Am Chem Soc* 132, 11779-11791, doi:10.1021/ja104903x (2010).

20 Georghiou, G., Kleiner, R. E., Pulkoski-Gross, M., Liu, D. R. & Seeliger, M. A. Development and structure-based mechanism of highly specific macrocyclic Src kinase inhibitors from a DNA-template library. *submitted* (2011).

21 Kim, Y. G., Lone, A. M., Nolte, W. M. & Saghatelian, A. Peptidomics approach to elucidate the proteolytic regulation of bioactive peptides. *Proc Natl Acad Sci U S A* 109, 8523-8527, doi:10.1073/pnas.1203195109 (2012).

22 Malito, E. *et al.* Molecular Bases for the Recognition of Short Peptide Substrates and Cysteine-Directed Modifications of Human Insulin-Degrading Enzyme†. *Biochemistry* 47, 12822-12834, doi:10.1021/bi801192h (2008).

23 Malito, E., Hulse, R. E. & Tang, W. J. Amyloid beta-degrading cryptidases: insulin degrading enzyme, presequence peptidase, and neprilysin. *Cellular and molecular life sciences : CMLS* 65, 2574-2585, doi:10.1007/s00018-008-8112-4 (2008).

24 Shen, Y., Joachimiak, A., Rosner, M. R. & Tang, W. J. Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. *Nature* 443, 870-874, doi:10.1038/nature05143 (2006).

25 Guo, Q., Manolopoulou, M., Bian, Y., Schilling, A. B. & Tang, W. J. Molecular basis for the recognition and cleavages of IGF-II, TGF-alpha, and amylin by human insulin-degrading enzyme. *J Mol Biol* 395, 430-443, doi:10.1016/j.jmb.2009.10.072 (2010).

26 Manolopoulou, M., Guo, Q., Malito, E., Schilling, A. B. & Tang, W. J. Molecular basis of catalytic chamber-assisted unfolding and cleavage of human insulin by human insulin-degrading enzyme. *J Biol Chem* 284, 14177-14188, doi:10.1074/jbc.M900068200 (2009).

27 Stella, V. J. & He, Q. Cyclodextrins. *Toxicologic pathology* 36, 30-42, doi:10.1177/0192623307310945 (2008).

28 Andrikopoulos, S., Blair, A. R., Deluca, N., Fam, B. C. & Proietto, J. Evaluating the glucose tolerance test in mice. *American journal of physiology. Endocrinology and metabolism* 295, E1323-1332, doi:10.1152/ajpendo.90617.2008 (2008).

29 Monzillo, L. U. & Hamdy, O. Evaluation of insulin sensitivity in clinical practice and in research settings. *Nutrition reviews* 61, 397-412 (2003).

30 Mirsky, I. A. & Perisutti, G. Effect of insulinase-inhibitor on hypoglycemic action of insulin. *Science* 122, 559-560 (1955).

31 Ahren, B., Winzell, M. S. & Pacini, G. The augmenting effect on insulin secretion by oral versus intravenous glucose is exaggerated by high-fat diet in mice. *The Journal of endocrinology* 197, 181-187, doi:10.1677/JOE-07-0460 (2008).

32 Winzell, M. S. & Ahren, B. The high-fat diet-fed mouse: a model for studying mechanisms and treatment of impaired glucose tolerance and type 2 diabetes. *Diabetes* 53 Suppl 3, S215-219 (2004).

33 Riddle, M. C. & Drucker, D. J. Emerging therapies mimicking the effects of amylin and glucagon-like peptide 1. *Diabetes Care* 29, 435-449 (2006).

34 Mooradian, A. D. & Thurman, J. E. Drug therapy of postprandial hyperglycaemia. *Drugs* 57, 19-29 (1999).

35 Malaisse, W. J. Pharmacology of the meglitinide analogs: new treatment options for type 2 diabetes mellitus. *Treatments in endocrinology* 2, 401-414 (2003).

36 Hollander, P. *et al.* Effect of pramlintide on weight in overweight and obese insulin-treated type 2 diabetes patients. *Obesity research* 12, 661-668, doi:10.1038/oby.2004.76 (2004).

37 Duckworth, W. C. & Kitabchi, A. E. Insulin and glucagon degradation by the same enzyme. *Diabetes* 23, 536-543 (1974).

38 Shroyer, L. A. & Varandani, P. T. Purification and characterization of a rat liver cytosol neutral thiol peptidase that degrades glucagon, insulin, and isolated insulin A and B chains. *Archives of biochemistry and biophysics* 236, 205-219 (1985).

39 Bennett, R. G., Duckworth, W. C. & Hamel, F. G. Degradation of amylin by insulin-degrading enzyme. *J Biol Chem* 275, 36621-36625, doi:10.1074/jbc.M006170200 (2000).

40 Authier, F., Mort, J. S., Bell, A. W., Posner, B. I. & Bergeron, J. J. Proteolysis of glucagon within hepatic endosomes by membrane-associated cathepsins B and D. *J Biol Chem* 270, 15798-15807 (1995).

41 Fontes, G. *et al.* Miniglucagon (MG)-generating endopeptidase, which processes glucagon into MG, is composed of N-arginine dibasic convertase and aminopeptidase B. *Endocrinology* 146, 702-712, doi:10.1210/en.2004-0853 (2005).

42 Trebbien, R. *et al.* Neutral endopeptidase 24.11 is important for the degradation of both endogenous and exogenous glucagon in anesthetized pigs. *American journal of physiology. Endocrinology and metabolism* 287, E431-438, doi:10.1152/ajpendo.00353.2003 (2004).

43 Young, A. Effects on plasma glucose and lactate. *Adv Pharmacol* 52, 193-208, doi:10.1016/S1054-3589(05)52010-6 (2005).

44 Gelling, R. W. *et al.* Lower blood glucose, hyperglucagonemia, and pancreatic alpha cell hyperplasia in glucagon receptor knockout mice. *Proc Natl Acad Sci U S A* 100, 1438-1443, doi:10.1073/pnas.0237106100 (2003).

45 Kolterman, O. G., Gottlieb, A., Moyses, C. & Colburn, W. Reduction of postprandial hyperglycemia in subjects with IDDM by intravenous infusion of AC137, a human amylin analogue. *Diabetes Care* 18, 1179-1182 (1995).

46 Riediger, T., Zuend, D., Becskei, C. & Lutz, T. A. The anorectic hormone amylin contributes to feeding-related changes of neuronal activity in key structures of the gut-brain axis. *American journal of physiology. Regulatory, integrative and comparative physiology* 286, R114-122, doi:10.1152/ajpregu.00333.2003 (2004).

47 Young, A. Inhibition of glucagon secretion. *Adv Pharmacol* 52, 151-171, doi:10.1016/S1054-3589(05)52008-8 (2005).

48 Schmitz, O., Brock, B. & Rungby, J. Amylin agonists: a novel approach in the treatment of diabetes. *Diabetes* 53 Suppl 3, S233-238 (2004).

49 Crowell, M. D., Mathis, C., Schettler, V. A., Yunus, T. & Lacy, B. E. The effects of tegaserod, a 5-HT receptor agonist, on gastric emptying in a murine model of diabetes mellitus. *Neurogastroenterology and motility : the official journal of the European Gastrointestinal Motility Society* 17, 738-743, doi:10.1111/j.1365-2982.2005.00681.x (2005).

50 Gedulin, B. R., Jodka, C. M., Herrmann, K. & Young, A. A. Role of endogenous amylin in glucagon secretion and gastric emptying in rats demonstrated with the selective antagonist, AC187. *Regulatory peptides* 137, 121-127, doi:10.1016/j.regpep.2006.06.004 (2006).

51 Unger, R. H. & Cherrington, A. D. Glucagonocentric restructuring of diabetes: a pathophysiologic and therapeutic makeover. *The Journal of clinical investigation* 122, 4-12, doi:10.1172/JCI60016 (2012).

52 Sadry, S. A. & Drucker, D. J. Emerging combinatorial hormone therapies for the treatment of obesity and T2DM. *Nature reviews. Endocrinology*, doi:10.1038/nrendo.2013.47 (2013).

Appendix 1

Metabolomics Experiments to identify *E.Coli* derived ligand of MR1

This Appendix is adapted from:

Jacinto, L et al. MAIT Recognition of a Stimulatory Bacterial Antigen Bound to MR1. The Journal of Immunology. 2013

A1.1 Introduction

MR1-restricted Mucosal Associated Invariant T (MAIT) cells represent a sub-population of αβ T cells with innate-like properties and limited TCR diversity. MAIT cells are of interest due to their reactivity against bacterial and yeast species, suggesting they play a role in defense against pathogenic microbes. Despite the advances in understanding MAIT cell biology, the molecular and structural basis behind their ability to detect MR1-antigen complexes is unclear. Here we present our structural and biochemical characterization of MAIT TCR engagement of MR1 presenting an *E. coli*-derived stimulatory ligand, rRL-6-CH_2OH previously found in *Salmonella typhimurium*. We show a clear enhancement of MAIT TCR binding to MR1 due to presentation of this ligand. Our structure of a MAIT TCR/MR1/rRL-6-CH_2OH complex shows an evolutionarily conserved binding orientation, with a clear role for both the CDR3α and CDR3β loops in recognition of the rRL-6-CH_2OH stimulatory ligand. We also present two additional xeno-reactive MAIT TCR/MR1 complexes that recapitulate the docking orientation documented previously despite having variation in the CDR2β and CDR3β loop sequences. Our data supports a model by which MAIT TCRs engage MR1 in a conserved fashion, their binding affinities modulated by the nature of the MR1 presented antigen or diversity introduced by alternate Vβ usage or CDR3β sequences.

A1.2 Results

These results prompted us to analyze by mass spectrometry the content of the hbMR1 in both *E. coli* exposed and non-exposed protein samples. Normal Phase QTOF analysis clearly identified a compound with a retention time of ~ 7 minutes present only in the *E. coli* treated sample (Figure 2A). Mass spectrometry analysis of this compound determined a compound with a m/z ratio of 329.1095 and whose fragmentation yielded a pattern nearly identical to that found for rRL-6-CH_2OH (r-RL) (Figure 2B), a MAIT cell stimulatory ligand characterized from *Salmonella typhimurium* supernatant[10]. This compound is generated as a by-product of the

riboflavin synthesis pathway and its structure comprises a lumazine core and a ribityl chain (Figure 2A, inset). rRL-6-CH_2OH has been shown to trigger MAIT cell activation in a MR1 dependent manner, and displayed the strongest potency in activating MAIT cells among three structurally related compounds[10]. Thus, the increased MAIT TCR affinity to the *E. coli*-treated hbMR1 sample correlates with the presence of this ribityl moiety, which our model suggested[9] directly participates in MAIT TCR recognition of MR1 presented antigen.

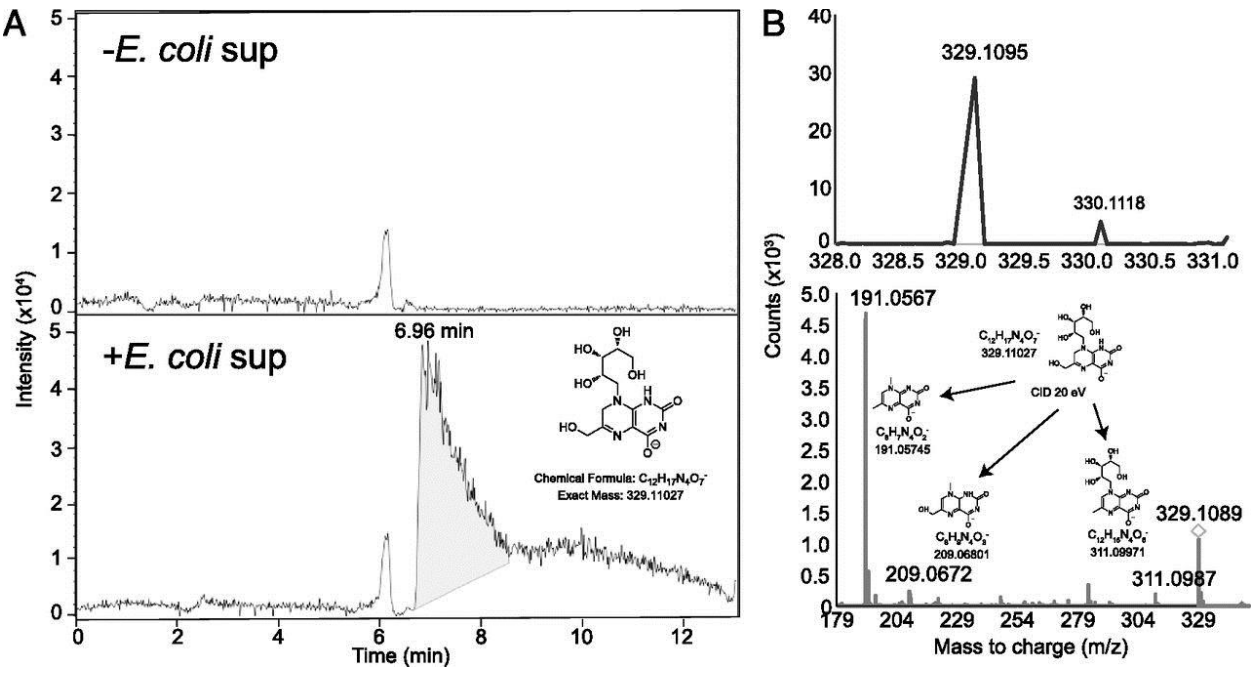


Figure A1.1 MAIT Recognition of a Stimulatory Bacterial Antigen Bound to MR1. MS reveals the presence of rRL-6-CH_2OH in the *E. coli*–treated hbMR1 sample. (**A**) Extracted ion chromatograms (EICs) for rRL-6-CH_2-OH (m/z 329.1103) from hbMR1 protein exposed to *E. coli* (*lower panel*) compared with untreated protein control (*upper panel*). EIC shows compound with m/z 329.1103 present only in *E. coli*–treated sample (*inset*). (**B**) Compound with m/z 329.1095 (left peak, *upper panel*) from the *E. coli*–treated hbMR1 sample and product ions from targeted fragmentation (*lower panel*); the structures of each of the products of the fragmentation are shown as *insets*. The precursor ion is indicated by a diamond. Tandem mass spectrometry (MS/MS) product ion data match against the theoretical fragmentation pattern of rRL-6-CH_2OH (right peak, *upper panel*) within <5 ppm.

A1.3 Methods

Mass spectrometry analysis of hbMR1. 6.25 µg of DMRL (m/z 325.1154) was injected onto a Luna NH2 4.6x50 mm, 5 µm column. The flow was analyzed by ESI-TOF-MS on a Bruker maXis impact QTOF with Agilent 1290 HPLC using a binary gradient of 95% B to 0% over 5 minutes (A: 20 mM ammonium acetate in water, pH 9.0; B: Acetonitrile). Target ion detected after elution with aqueous gradient; data was collected in negative ion mode. Retention time was obtained by extracted ion chromatogram of respective m/z; product ions were obtained by tandem MS with a collision energy of 20eV. For hbMR1, comparison of hbMR1 protein that was exposed to *E. coli* sn and an untreated control hbMR1 sample were analyzed on a Bruker maXis impact QTOF LC/MS in negative ion mode. 5 uL each of 25 µM hbMR1 was injected onto a Luna NH2 4.6x50 mm, 5 um column in 20 mM ammonium acetate, pH 9 buffer and eluted with an aqueous gradient as described above. Retention time of rRL-6-CH_2OH was determined by extracted ion chromatograms using the m/z values. Product ions were obtained from target fragmentation at collision induced voltages of 20.

Appendix 2 Evaluation of Potent α/β -Peptide Analogue of GLP-1 with Prolonged Action In Vivo

Adapted from prepared manuscript:

Johnson, LM et al. Proteolysis in the Context of Diabetes: Evaluation of Potent α/β-Peptide Analogue of GLP-1 with Prolonged Action In Vivo. 2014

A2.1 Introduction

Glucagon-like peptide-1 (GLP-1) is the natural agonist for GLP-1R, a G protein-coupled receptor (GPCR) that is displayed on the surface of pancreatic β cells. Activation of GLP-1R augments glucose-dependent insulin release from β cells and promotes β cell survival. These properties are attractive for treatment of type 2 diabetes, but GLP-1 is rapidly degraded by peptidases *in vivo*. Efforts to develop small-molecule agonists of GLP-1R have not been successful, presumably because receptor activation requires contact over an extended surface. All potent GLP-1R agonists reported to date are comprised exclusively of α-amino acid residues, with prolonged *in vivo* activity achieved by variations in sequence, incorporation of stabilizing appendages, and/or specialized delivery strategies. We describe an alternative design strategy for retaining GLP-1-like function but engendering prolonged activity *in vivo*, based on strategic replacement of native α residues with conformationally constrained β-amino acid residues. This approach may be generally useful for developing stabilized analogues of peptide hormones. As part of this collaboration, with the Gellman laboratory at University of Wisconsin, Madison, I ran comparative glucose tolerance tests in mice to assay the efficacy of different GLP-1 analogues.

A2.2 Results

The ability of α/β -peptide **6** to augment insulin secretion from mouse pancreatic islets in response to elevated glucose led us to evaluate the activity of this GLP-1 analogue *in vivo* via glucose tolerance tests (GTT) (Figure 3). In addition to the α/β -peptide and GLP-1(7-37)-NH_2, these studies included exendin-4 (39 residues), all three of which were tested for the ability to normalize circulating glucose levels. GLP-1, exendin-4 and α/β -peptide **6** were compared at a dose of 1 mg/kg, and descending doses were examined for the α/β -peptide. For mice injected with vehicle rather than peptide (negative control), the intraperitoneal glucose challenge causes a rapid rise in blood glucose concentration that subsides after 30 min (Figure 3A-B). Mice

injected with GLP-1, exendin-4 or α/β -peptide **6** at 1 mg/kg showed a dramatic suppression in the rise of glucose relative to vehicle-treated mice during the GTT; the three compounds were equally effective at this dose. Dose-response behavior was observed for **6**, with glucose control maintained at 0.1 mg/kg, but not at 0.01 mg/kg.

In order to determine whether the GLP-1R agonists display prolonged action, I repeated the GTT 5 hr after agonist administration (Figure 3C). Mice treated with GLP-1(7-37)-NH_2 showed no significant difference from those treated with vehicle 30 min after the second glucose challenge; this result is expected based on the rapid enzymatic inactivation of GLP-1 *in vivo*. In contrast, the glucose-lowering effect of exendin-4 is maintained at 5 hr, which is consistent with prior observations. Exendin-4 persists longer in the bloodstream than does GLP-1 because exendin-4 is not cleaved by DPP-4 and is only very slowly degraded by neprilysin or other peptidases.[12,13] α/β -Peptide **6** exerted the same glucose-lowering effect at 5 hr as exendin-4 (each at 1 mg/kg). The ability of **6** to induce control of blood glucose levels at 5 hr was manifested even when the α/β -peptide was administered at 0.1 mg/ml. These results presumably reflect the resistance of this α/β -peptide to degradation by either DPP-4 or neprilysin.

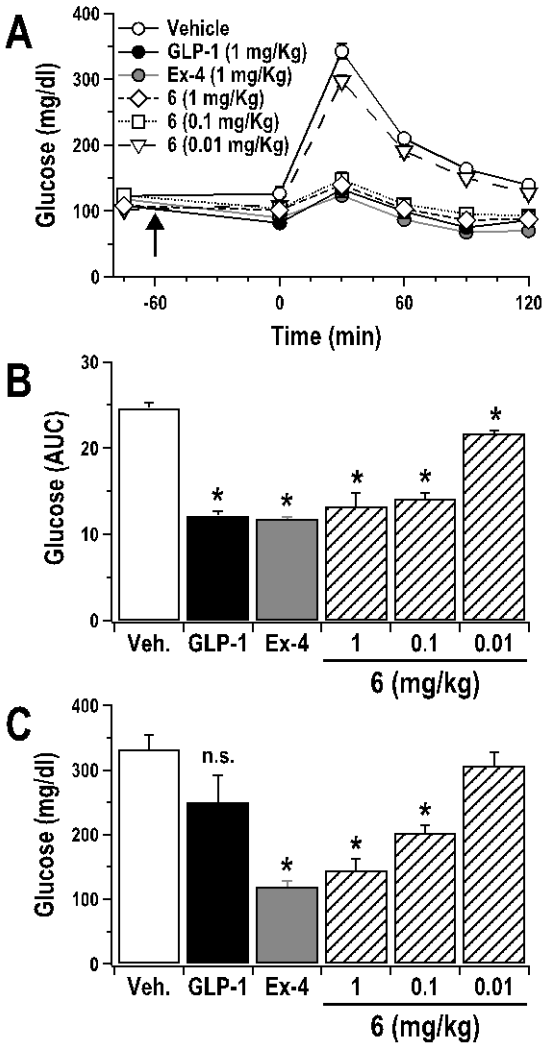


Figure A2.1 α/β-peptide 6 demonstrates long lasting improvement of *in vivo* glucose tolerance.

A) Plasma glucose values during a glucose tolerance test (GTT) for mice treated with GLP-1(7-37)-NH_2 (1 mg/kg), Exendin-4 (Ex-4, 1 mg/kg), varying doses (0.01-1 mg/kg) of α/β-peptide **6** or vehicle. Upward arrow indicates timing of the peptide treatments delivered via IP injection. Results show mean (± SEM) of 4 separate mice per condition. **B**) Average area under the curve (AUC) values for the GTT data shown in **A**. **C**) Plasma glucose values at 30 mins following a second GTT that was conducted 5 hr following that shown in **A**. *, $P < 0.05$ vs. vehicle.

A2.3 Discussion

Our results show that a GLP-1-derived oligomer containing multiple replacements of □-amino acid residues with conformationally constrained β-amino acid residues can maintain native-like agonist activity at the GLP-1 receptor. α/β -Peptide **6** mimics GLP-1 in terms of augmenting glucose-stimulated insulin secretion from pancreatic β cells and regulating blood glucose levels *in vivo*. The prolonged effect of this α/β -peptide relative to GLP-1 is attributed to the strongly diminished susceptibility to degradation by widely distributed peptidases, a property that arises in part from the multiple β residue replacements. The favorable characteristics displayed by α/β -peptide **6** *in vitro* and *in vivo* suggest that backbone-modification strategies involving periodic replacement of α residues with β residues may prove to be generally useful for developing potent and highly stable analogues of peptide hormones and other signaling peptides for which the receptor-bound conformation is at least partially α-helical. Side chain cross-links have been widely employed for this purpose in the past,[17-29] but recent structural studies have revealed that the large additional molecular surfaces generated by the cross-link itself can make intimate contacts with partner protein surfaces.[44-46] Such contacts might alter binding affinity and/or selectivity relative to the natural prototype. Periodic α-to-β replacement may represent a more effective approach to retaining an α-helix-like shape[47] and the binding selectivity of a prototype α-peptide while diminishing susceptibility to proteolytic action.

A2.4 Method

Glucose Tolerance Test (GTT). GLP-1(7-37)-NH_2, exendin-4 or α/β-peptide **6** was administered to mice by interperitoneal (i.p.) injection at a 1 mg/kg dose, or lower for α/β-peptide **6**, using an injection volume of 10 mL/kg body mass. Each peptide was first dissolved in prefiltered DMSO at 10 mg/mL concentration, then diluted >20-fold with TBS buffer, pH 7.4 (final DMSO conc. = <5%). Glucose was administered by i.p. injection with a sterile-filtered 30% D-glucose-saline solution at a 1.5 g/kg dose using a 5 mL/kg injection volume.

Thirteen-week-old male C57BL/6J mice (n = 4) were fasted overnight on wood chip bedding for 15 hours prior to the experiment. Blood glucose levels were monitored from a tail tip bleed using an ACCU-CHEK Aviva blood glucose meter. Fasting glucose levels were measured at 75 minutes prior to the glucose injection (t = -75 min), and the compound injection was performed 60 minutes prior to the glucose injection (t = -60 min). Glucose levels were measured immediately prior to the glucose injection (t = 0 min) to assess any changes in the baseline glucose caused by peptide administration. Blood glucose levels were monitored at 30, 60, 90, and 120 minutes after injection of glucose.

Five hours after peptide injection, the mice received a second injection of 1.5 g/kg glucose. Blood glucose was monitored at 30 min after this second glucose injection. Mice were sacrificed by CO_2 inhalation at the conclusion of the GTT.

A2.5 References

12. Gao, W. & Jusko, W.J. Target-mediated pharmacokinetic and pharmacodynamic model of exendin-4 in rats, monkeys, and humans. *Drug Metab. Dispos.* **40**, 990-7, (2012).

13. Parkes, D., Jodka, C., Smith, P., Nayak, S., Rinehart, L., Gingerich, R., Chen, K. & Young, A. Pharmacokinetic actions of exendin-4 in the rat: Comparison with glucagon-like peptide-1, *Drug Dev. Res.* **53***,* 260-7, (2001).

17. Felix, A.M., Heimer, E.P., Wang, G.T., Lambros, T.J., Fournier, A.J., Mowles, T.F., Maines, S., Campbell, R.M., Wegrzynski, B.B., Toomer, V., Fry, D., & Madison, V.S. Synthesis, biological activity and conformational analysis of cyclic GRF analogs. *Int. J. Peptide Protein Res.* **21**, 441-54, (1988).

18. Ghadiri, M.R., & Choi, C. Secondary structure nucleation in peptides. Transition metal ion stabilized □-helices. *J. Am. Chem. Soc.* **112**, 1630-32, (1990).

19. Jackson, D.Y., King, D.S., Chmielewski, J., Singh, S., & Schultz, P.G. General approach to the synthesis of short α-helical peptides. *J. Am. Chem. Soc.* **113**, 9391-2, (1991).

20. Chorev, M., Roubini, E., McKee, R.L, Gibbons, S.W., Goldman, M.E., Caufield, M.P., & Rosenblatt, M. Cyclic parathyroid hormone related protein agonists: lysine 13 to aspartic acid 17 [*i* to (*i+4*)] side chain to side chain lactamization. *Biochemistry* **30**, 5968-74, (1991).

21. Judice, J.K., Tom, J.Y.K., Huang, W., Wrin, T., Vennari, J., Petropoulos, C.J., & McDowell, R.S. Inhibition of HIV type 1 infectivity by constrained □-helical peptides: implications for the viral fusion mechanism. *Proc. Natl. Acad. Sci. USA* **94**, 13426-30, (1997).

22. Trivedi, D., Lin, Y., Ahn, J.M., Siegel, M., Mollova, N.N., Schram, K.H., & Hruby, V.J. Design and synthesis of conformationally constrained glucagon analogues. *J. Med. Chem.* **43**, 1714-22, (2000).

23. Blackwell, H.E. & Grubbs, R.H. Highly efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis. *Angew. Chem. Int. Ed.* **37**, 3281-4, (1998).

24. Walensky, L.D., Kung, AL., Escher, I., Malia, T. J., Barbuto, S., Wright, R.D., Wagner, G., Verdine, G.L. & Korsmeyer, S.J. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* **305**, 1466-70, (2004).

25. Chapman, R.N., Dimartino, G. & Arora, P.S. A highly stable short □-helix constrained by main-chain hydrogen-bond surrogate. *J. Am. Chem. Soc.* **126,** 12252-3, (2004).

26. Verdine, G.L. & Hilinski, G.J. Stapled peptides for intracellular drug targets. *Meth. Enzymol.* **503**, 3-33, (2012).

27. Okamoto, T., Zobel, K., Fedorova, A., Quan, C., Yang, H., Fairbrother, W.J., Huang, D.C. S., Smith, B.J., Keshayes, K. & Czabotar, P.E. Stabilization the pro-apoptotic BimBH3 helix (BimSAHB) does not necessarily enhance affinity of biological activity. *ACS Chem. Biol.* **8**, 297-302, (2013).

28. Miranda, L.P., Winters, K.A., Gegg, C.V., Patel, A., Aral, J., Long, J., Zhang, J., Diamond, S., Guido, M., Stanislaus, S., Ma, M., Li, H., Rose, M.J., Poppe, L., & Veniant, M.M. Design and synthesis of conformationally constrained glucagon-like peptide-1 derivatives with increased plasma stability and prolonged in vivo activity. *J. Med. Chem.* **51**, 2758-65, (2008).

29. Murage, E.N., Gao, G.Z., Bisello, A., & Ahn, J.M. Development of Potent Glucagon-like Peptide-1 Agonists with High Enzyme Stability via Introduction of Multiple Lactam Bridges. *J. Med. Chem. 53*, 6412-20, (2010).

44. Stewart, M.L., Fire, E., Keating, A.E., & Walensky, L.D. The Mcl-1 BH3 helix is an exclusive Mcl-1 inhibitor and apoptosis sensitizer, *Nat. Chem. Biol. **6***, 595-601, (2010).

45. Phillips, C., Roberts, L.R. Schade, M., Bazin, R., Bent, A., Davies, N.L., Moore, R., Pannifer, A.D., Pickford, A.R., Prior, S.H., Read, C.M., Scott, A., Brown, D.G., Xu, B., & Irving, S.L.,

Design and structure of stapled peptides binding to estrogen receptors, *J. Am. Chem. Soc.* **133,** 9696-9, (2011).

46. Baek, S., Kutchukian, P.S., Verdine, G.L., Huber, R., Holak, T.A., Lee, K.W., & Popowicz, G.M., Structure of the stapled p53 peptide bound to Mdm2, *J. Am. Chem. Soc.* **134**, 103-6, (2012).

47. Bullock, B.N., Jochim, A.L., & Arora, P.S. Assessing helical protein interfaces for inhibitor design. *J. Am. Chem. Soc.* **133**, 14220-3, (2011).

Appendix Chapter 3: Supplementary Data

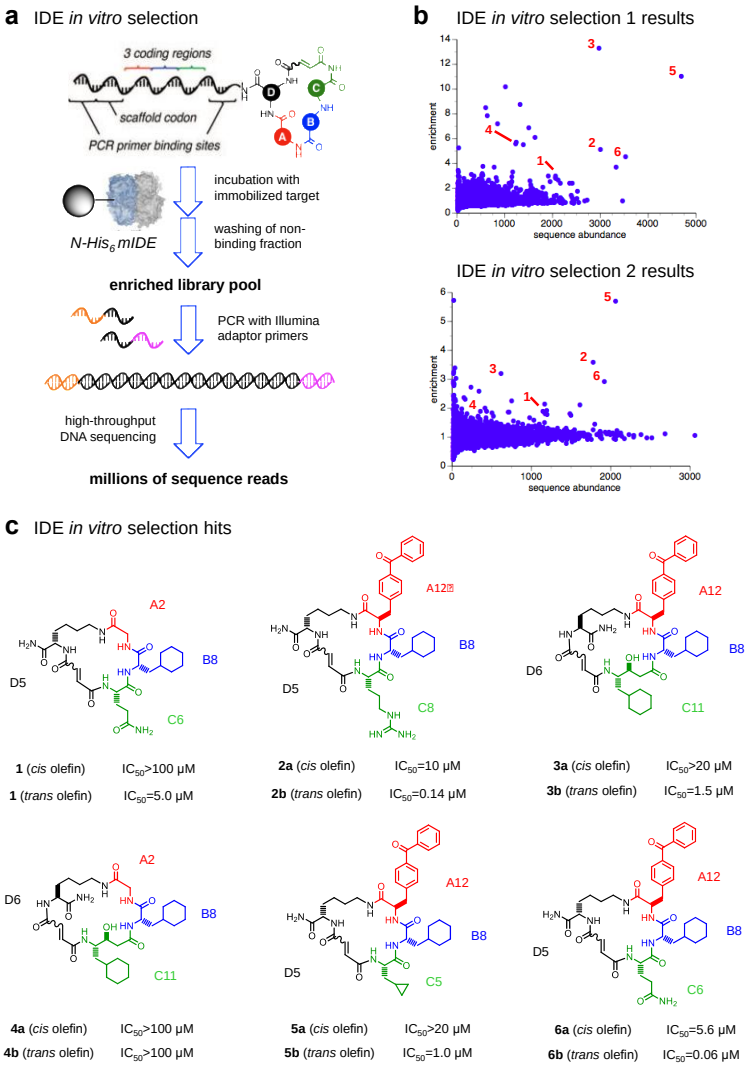


Figure A3.1 Inhibitor Selection Scheme a, Overview of the *in vitro* selection of a 13,824-membered DNA-templated macrocycle library for IDE binding affinity[18,19]. **b**, Enrichment results from two independent *in vitro* selections against N-His_6-mIDE using the DNA-templated macrocycle library[18]. The numbers highlight compounds enriched at least 2-fold in both selections. **c**, Structures of IDE-binding macrocycles **1**-**6** decoded from DNA library barcodes corresponding to building blocks A, B, C and D. The *cis* and *trans* isomers are labeled '*a*' and '*b*', respectively. The two isomers were synthesized as previously reported[18,19] and separated by HPLC. IDE inhibition activity was assayed by following cleavage of the fluorogenic peptide substrate Mca-RPPGFSAFK(Dnp)-OH.

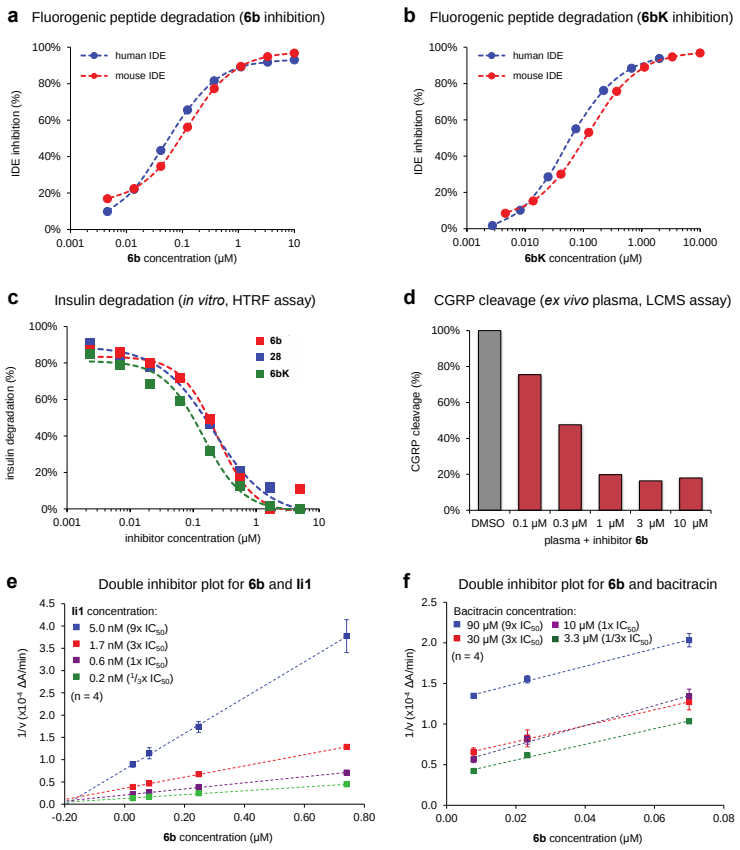


Figure A3.2 Inhibition of human and mouse IDE activity demonstrated using distinct assays. **a** and **b**, cleavage of the fluorogenic substrate peptide Mca-RPPGFSAFK(Dnp)-OH by human and mouse IDE in the presence of inhibitors (**a**) **6b** and (**b**) **6bK**. **c**, Homogeneous time-resolved fluorescence (HTRF, Cisbio) assay measuring degradation of insulin by IDE (R&D) in the presence of **6b**, **6bK** and **28**. **d**, LC-MS assay for *ex vivo* degradation of CGRP (10 μM) by endogenous IDE in mouse plasma in the presence of **6b**. **e** and **f**, Biochemical assays suggesting that **6b** binds a site in IDE distinct from the conventional peptide substrate binding site known to bind substrate mimetic **Ii1**. Yonetani–Theorell double inhibitor plots of IDE activity in the presence of (**e**) **6b** and **Ii1**, or (**f**) **6b** and bacitracin. Crossing lines indicate synergistic and independent binding of inhibitors, while parallel lines indicate competition for binding to the enzyme. The inhibitors assayed using the fluorogenic substrate Mca-RPPGFSAFK(Dnp)-OH.

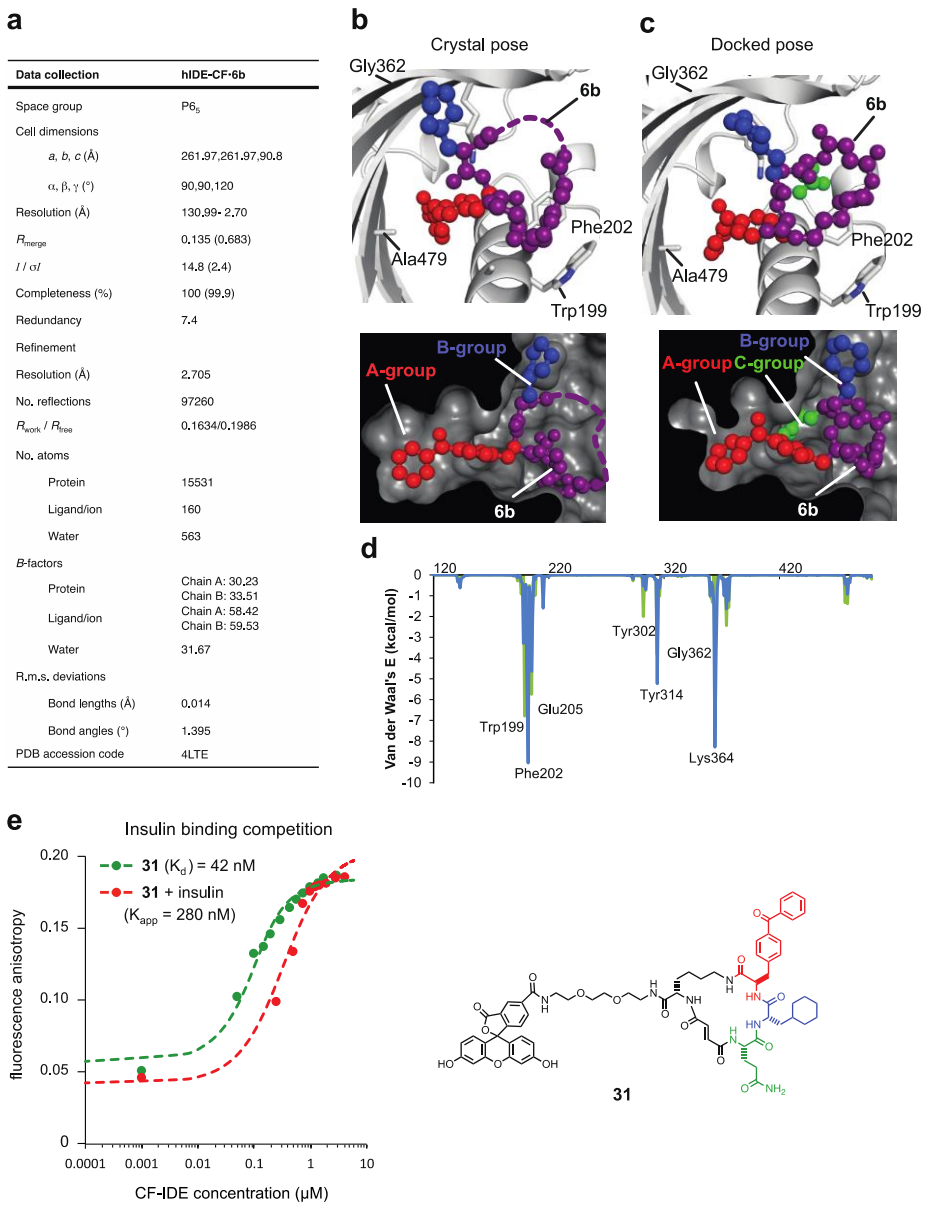


Figure A3.3 Data collection and refinement statistics (molecular replacement), docking simulation for 6b, and competition test between insulin and fluorescein-labeled macrocycle 31 for binding CF-IDE. **a**, One crystal was used to solve the CF-IDE•**6b** structure. The highest-resolution shell is shown in parentheses. Structure coordinates are deposited in the Protein Data Bank (accession number 4TLE). **b**, Molecular docking simulations are consistent with the placement of building blocks A and B in the structural model. The structure

of **6b** in binding site from crystallographic data with composite omit map contoured at 1.0 σ (p-benzoyl-phenylalanine is shown in red, cyclohexylalanine in blue, and the backbone in purple). **c,** Highest-scoring pose from DOCK simulations (glutamine group is shown in green). **d**, Residue decomposed energy of the crystal (green) and docked (blue) poses of **6b**. **e**, Competition test between the fluorescein-labeled macrocycle **31** and insulin for binding CF-IDE. Cysteine-free catalytically-inactive IDE was titrated against 0.9 nM macrocycle **31** alone (●) or together with 2.15 μM insulin (●), producing a shift in apparent dissociation constant for macrocycle **31** according to equation **Eq. 1** (Supplementary Information). Representative fluorescence polarization plots are shown.

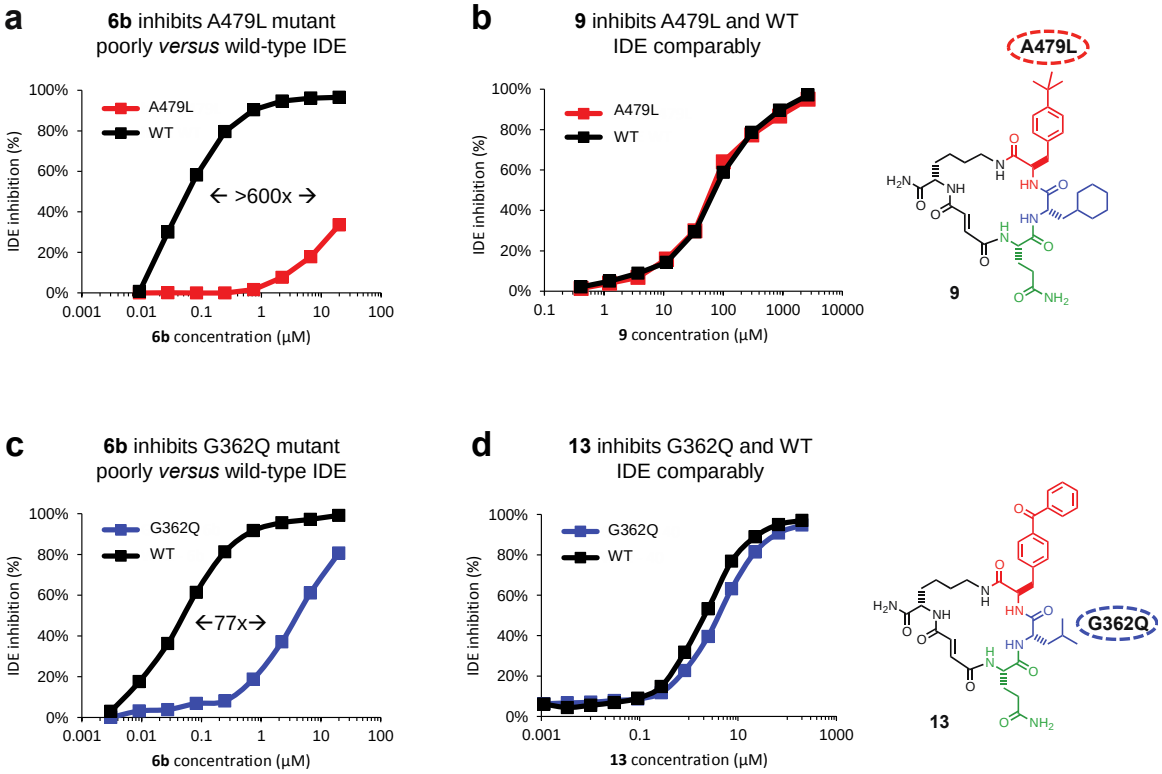


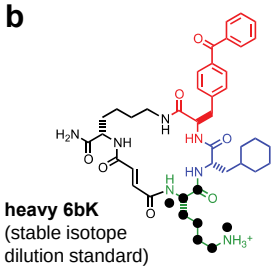
Figure A3.4 Small molecule-enzyme mutant complementation study to confirm the macrocycle binding site and placement of the benzophenone and cyclohexyl building-block groups. **a**, IDE mutant A479L is inhibited by **6b** >600-fold less potently compared to wild-type IDE. **b**, Analog **9**, in which the *p*-benzoyl ring is substituted for a smaller *tert*-butyl group, inhibits A479L IDE and WT IDE comparably. **c**, Similarly, IDE mutant G362Q is inhibited 77-fold less potently by **6b** compared with WT IDE. **d**, Analog **13**, in which the L-cyclohexyl alanine side chain was substituted with a smaller L-leucine side chain, inhibits G362Q IDE and WT IDE comparably. The full list of IDE mutants investigated is shown in Supplementary Table S5.

a

6bK stability *ex vivo*

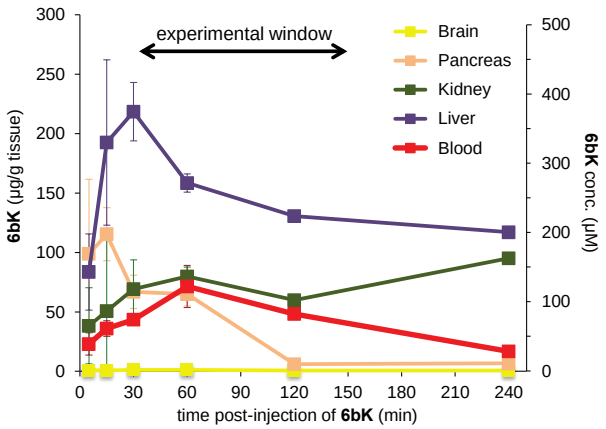
Sample	Plasma binding	Plasma stability (1 h)	Microsomal stability (1 h)
mouse	94 %	74 %	78 %
human	86 %	83 %	80 %

b



heavy 6bK
(stable isotope
dilution standard)

c



d

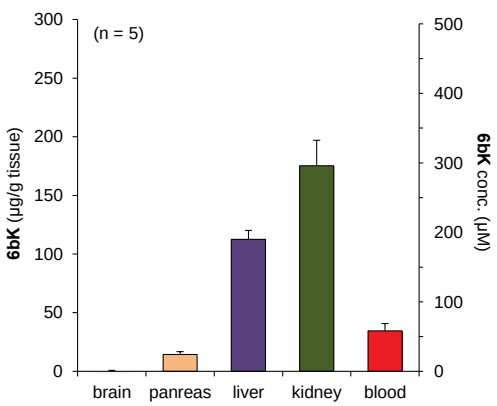


Figure A3.5 Pharmacokinetic parameters of 6bK. **a,** Plasma binding, plasma stability, and microsomal stability (1 h incubation) data was provided by Dr. Stephen Johnston and Dr. Carrie Mosher (Broad Institute). **b**, Heavy **6bK** was synthesized with ^{15}N,^{13}C-lysine for stable-isotope dilution LC-MS quantitation. **c** Concentration of **6bK** in mice tissues and plasma collected over 4 hours (n = 1-2). **d**, Average biodistribution of **6bK** in five lean mice at 150 min post-injection of **6bK** 80 mg/kg i.p. at the endpoint of a IPGTT experiment. We did not detect **6bK** in the brain even using 10-fold concentrated samples for LC-MS injection compared to other tissues.

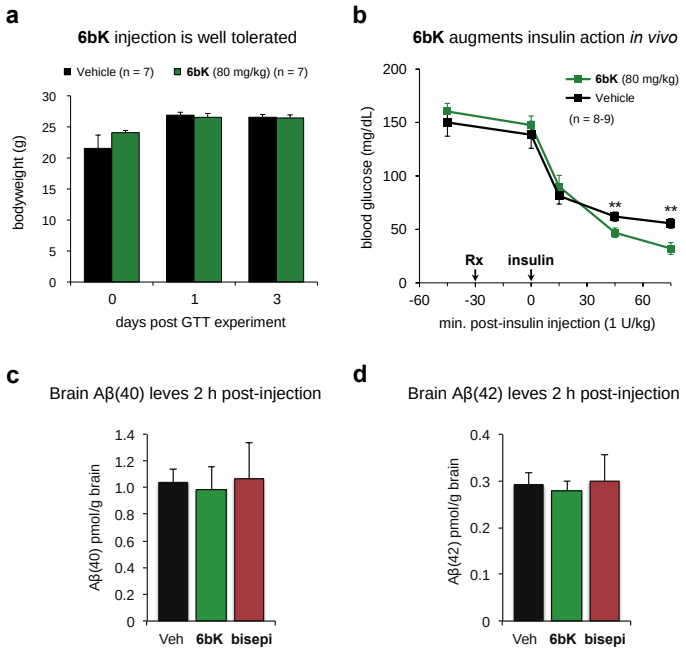


Figure A3.6 Dose tolerance, augmented insulin hypoglycemic action by 6bK in mice, and unaffected amyloid peptide levels in the mouse brain. **a**, Body weight measurements for C57BL/6J mice treated with **6bK** (■, 80 mg/kg i.p.) or vehicle alone (■). Mice treated with **6bK** (■, 80 mg/kg) are active, display normal posture, normal grooming, and response to stimulation. **b**, Increased hypoglycemic response to a high dose of insulin (1.0 U/kg) under acute IDE inhibition (compare with Fig. 4b). Non-fasted male C57BL/6J mice were treated with a single i.p. injection of IDE inhibitor **6bK** (■, 80 mg/kg) or vehicle alone (■), followed by i.p. insulin (Humulin-R, 1.0 U/kg). The animals in the **6bK**-treated cohort experienced lethargy and hypothermia consistent with the augmented effect of insulin. **c** and **d**, Treatment of C57BL/6J lean mice with **6bK** (■, 80 mg/kg, n = 6) does not change brain levels of Aβ(40) or Aβ(42) peptides in the brain 2 h post injection compared to treatment with vehicle alone (■, n = 5) or inactive diastereomer **bisepi-6bK** (■, 80 mg/kg, n = 6). All data points and error bars represent mean ± SEM.

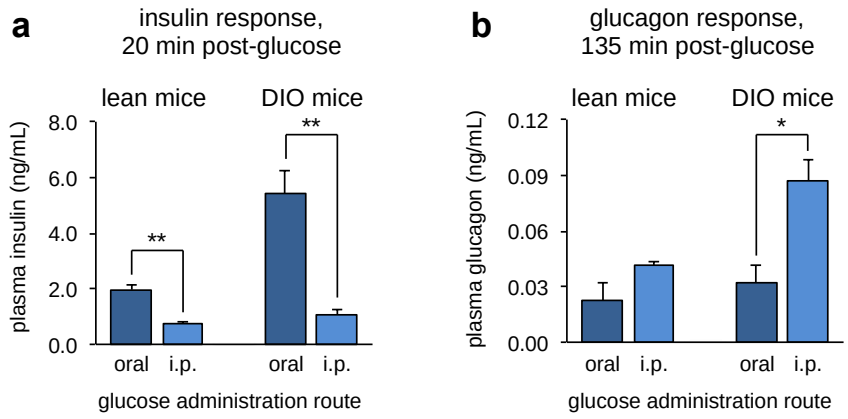


Figure A3.7 Dependence of insulin and glucagon secretion on the route of glucose administration (oral or i.p.) due to the both the 'incretin effect' as well as the hyperinsulinemic phenotype of DIO *versus* lean mice. **a**, The early insulin response to glucose in lean and DIO mice is higher during OGTT than IPGTT. **b**, Suppression of glucagon secretion post-glucose administration is less effective after IPGTT and in DIO mice. All data points and error bars represent mean ± SEM. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures are * $p < 0.05$ or ** $p < 0.01$ between the labeled groups.

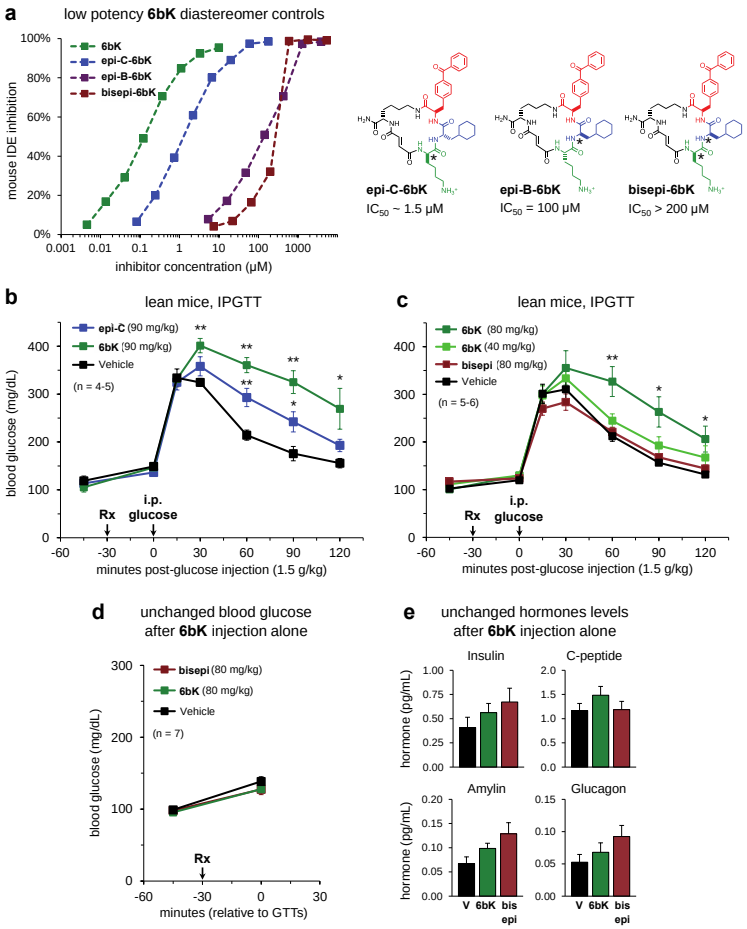


Figure A3.8 Low-potency diastereomers of 6bK used to determine effective dose range of 2 mg/mouse and confirm on-target IDE inhibition effects during IPGTTs in lean and DIO mice. **a**, Inhibition of mouse IDE activity by low potency diastereomers of **6bK**. The stereocenters altered in each compound relative to those of **6bK** are labeled with a star. **b**, The effects of **6bK** (■, 90 mg/kg) were compared to the weakly active stereoisomer **epi-C-6bK** (■, 90 mg/kg) and vehicle controls (■) during an ipGTT to determine the dosing range in lean mice. **c**, Two doses of **6bK**, 80 mg/kg (■) and 40 mg/kg (■) were compared with inactive **bisepi-6bK** (■, 80 mg/kg) and vehicle alone (■). **d** and **e**, Administration of **6bK** (■, 80 mg/kg) to lean mice not followed by injection of a nutrient such as glucose or pyruvate did not significantly alter (**d**) basal blood glucose or (**e**) basal hormone levels compared to **bisepi-6bK** (■) or vehicle controls

(■). All data points and error bars represent mean ± SEM. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures are * $p < 0.05$ versus vehicle control group; ** $p < 0.01$ versus vehicle control group.

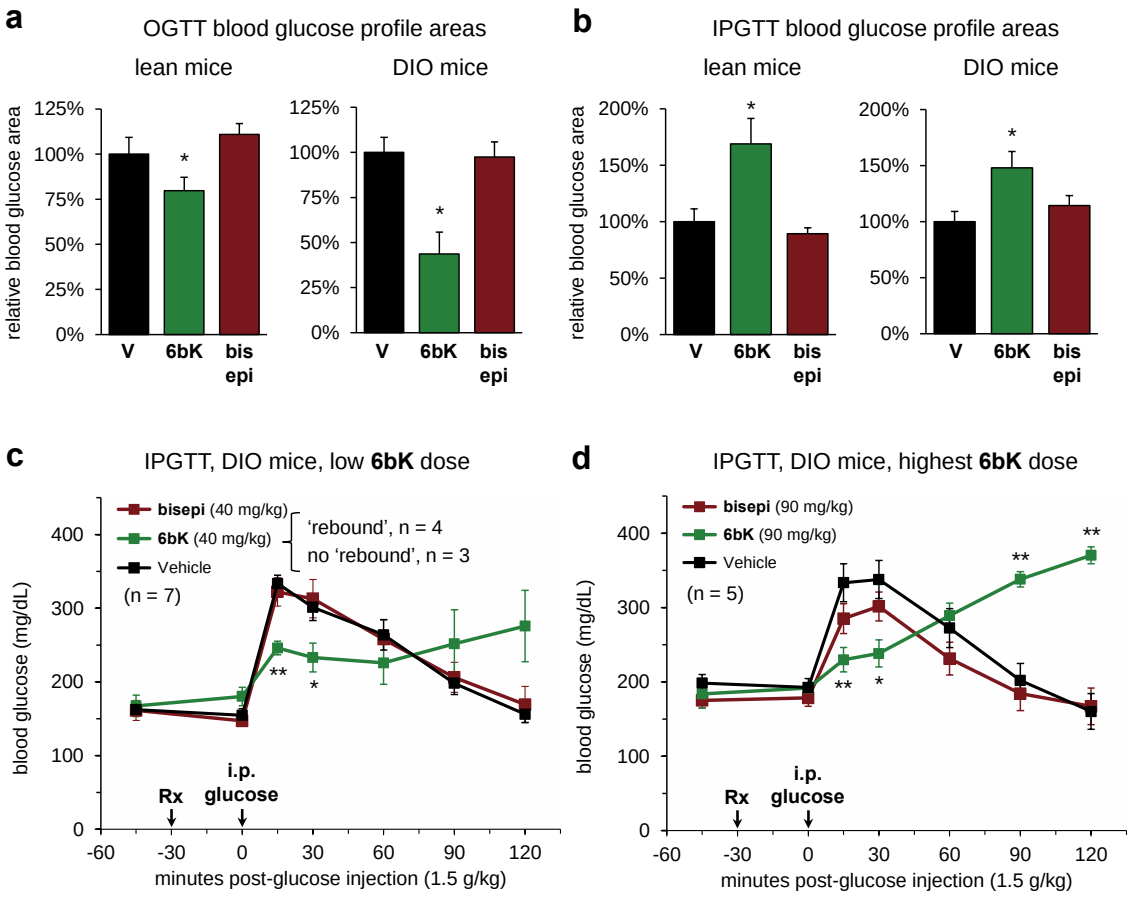


Figure A3.9 Relative area under the curve calculations and 6bK dose response for the glucose tolerance tests shown in Fig. 3.3. **a**, During OGTT, lean and DIO mice treated with **6bK** (■, 2 mg/animal) display improved glucose tolerance (lower blood glucose area), compared to vehicle controls (■) and inactive **bisepi-6bK** (■). **b**, In contrast, during IPGTT both lean and DIO mice treated with **6bK** (■) display impaired glucose tolerance (higher blood glucose area) compared to vehicle (■) or **bisepi-6bK** (■) controls. **c** and **d**, Dose-response of **6bK** (40 and 90 mg/kg; see Fig. 3d for 60 mg/kg) followed by IPGTT in DIO mice. Vehicle alone (■) and a matching dose of **bisepi-6bK** (■) were used as controls. **c**, DIO mice treated with low doses of **6bK** (■, 40 mg/kg) responded to IPGTT in either of two ways: improved glucose tolerance throughout the experiment (n = 3) or a hyperglycemic rebound as described in the main text (n =

4), suggesting this dose is too low to achieve a consistent effect (note the large error bars). **d**, Mice treated with high doses of **6bK** (■, 90 mg/kg) respond similarly to 60 mg/kg (Fig. 3d), but the potential weak activity observed for **bisepi-6bK** (IC_{50} > 100 μM) at this high dose (■) compared to vehicle alone (■) suggests that 90 mg/kg may be too high. All data points and error bars represent mean ± SEM. Statistics were performed using a two-tail Student's t-test, and significance levels shown in the figures is * $p < 0.05$ versus vehicle control group. See the Supplementary Methods section for a description of the AUC calculation.

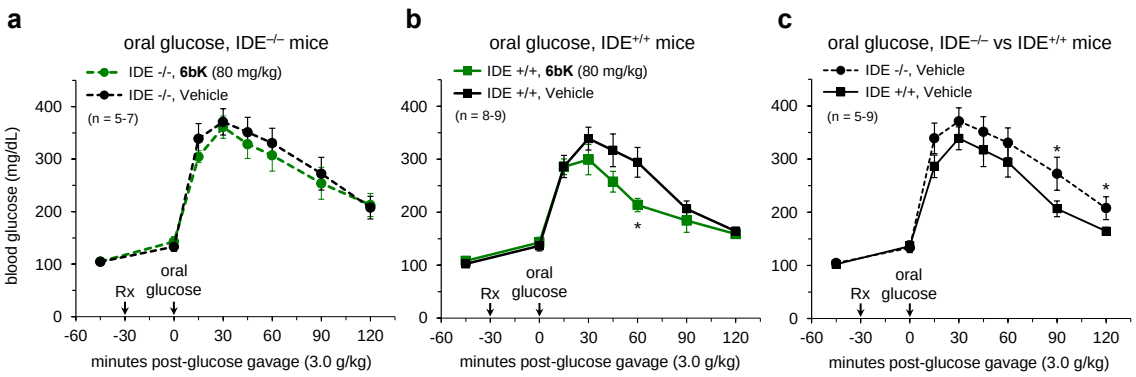
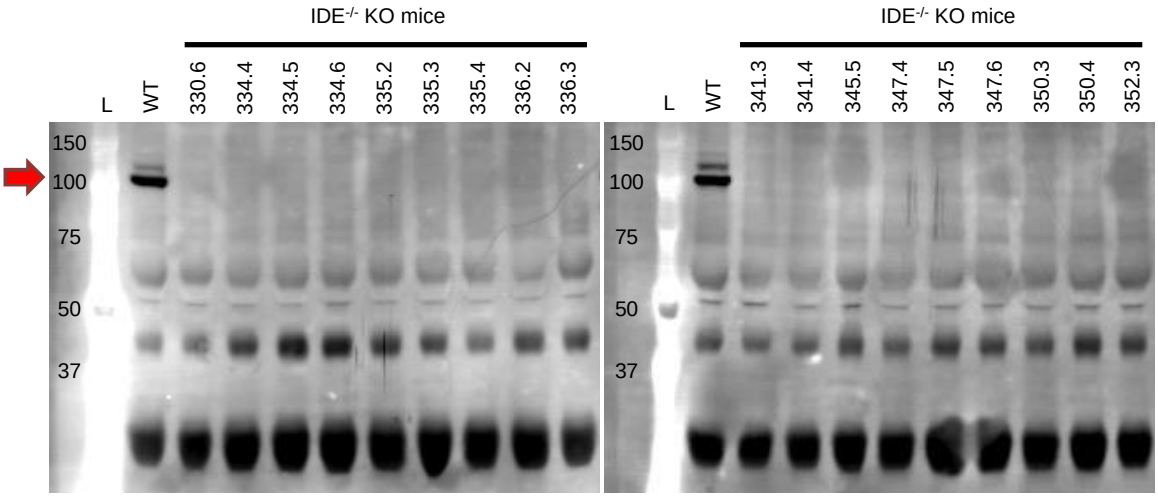
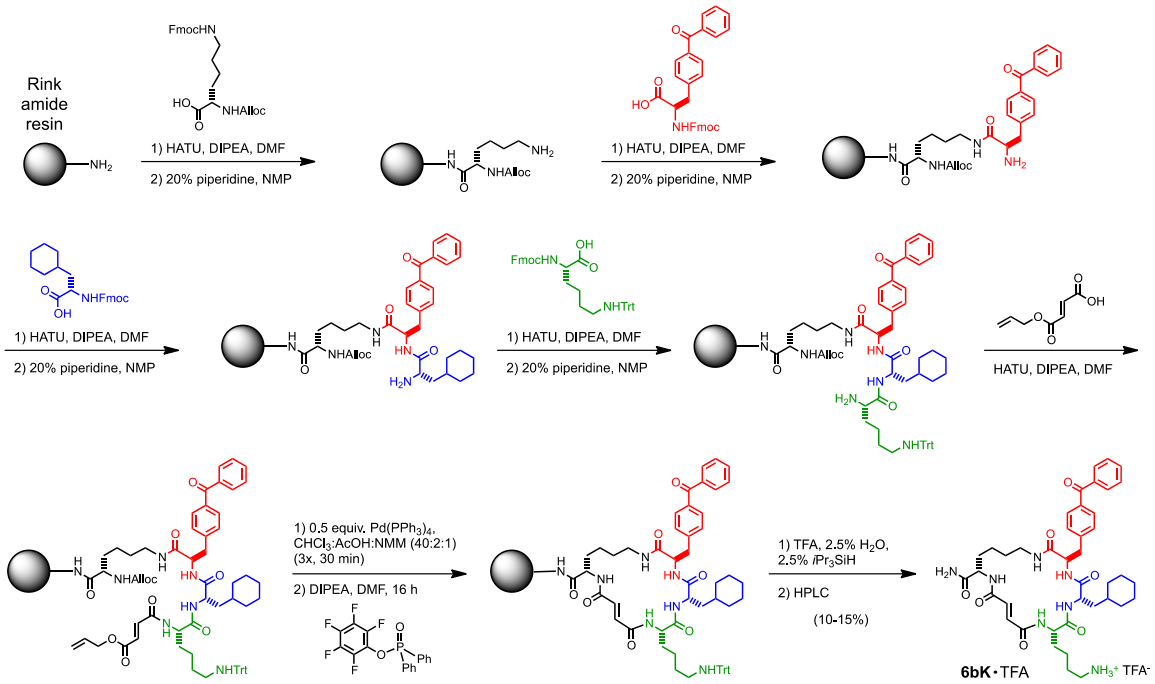


Figure A3.10 Oral glucose tolerance of mice lacking IDE compared to WT lean mice, and lack of effect of 6bK treatment in IDE$^{-/-}$ knockout mice. a, Mice lacking IDE treated with **6bK** (●, 80 mg/kg, i.p.) followed by oral glucose produce an identical response compared to vehicle-treated controls (●) of the same genotype (IDE$^{-/-}$). **b**, Transient IDE inhibition in age-matched IDE$^{+/+}$ WT control mice treated with **6bK** (■, 80 mg/kg, i.p.) display improved oral glucose tolerance compared to vehicle-treated controls (■) of the same genotype (similar to Fig. 3a). **c**, In contrast, when mice lacking IDE (●, IDE$^{-/-}$) are compared with age-matched IDE$^{+/+}$ controls (■) the IDE knockout mice display *impaired* oral glucose tolerance, supporting the hypothesis that chronic deletion of IDE leads to pronounced compensatory changes in metabolism that are not predictive of the outcome of IDE inhibition.[10,11] All data points and error bars represent mean ± SEM. Statistics were performed using a two-tail Student's t-test, and significance level shown in the figures is * $p < 0.05$ versus WT vehicle control group.



Supplementary Figure S1. Western blot for IDE using whole blood lysate (0.5 µL/well) from IDE$^{-/-}$ mice and a wild-type (WT) control sample.



Supplementary Scheme S1. Fmoc-based solid-supported synthesis of 6bK followed by on-resin macrocyclization.

IDE substrates *in vitro* (reference)	Alternative degrading enzymes (Brenda db)	k_{cat} (min^{-1})	K_M (μM)	k_{cat}/K_M	IC_{50} (μM)	X-ray struct.	IDE substrate *in vivo* (ref.)
insulin [28,29]	lysosomal proteases, disulfide reductase	1.52	<0.03	50	-	2WBY 2WC0	KO study [30,31]
Aβ (40, 42) [32,33]	NEP, cathepsin D	52	1.23	43	-	2G47 2WK3	KO study [30,34,35]
calcitonin-gene related peptide [6]	ECE, PREP	-	-	-	-	-	KO study [6]
glucagon [36,37]	NEP, nardilysin, DPP-IV, cathepsins B and C	38.5	3.46	11	weak	2G49	this study
amylin [38]	-	n/a	*0.3**	n/a	0.16	3HGZ 2G48	this study
TGF-α [39]	-	-	*0.15**	-	0.08	3E50	-
IGF-1 [34,40]	DPP-IV	-	weak	-	-	-	-
IGF-2 [34,40]	-	-	*0.1**	-	0.06	3E4Z	-
somatostatin-14 [41]	nardilysin, dactylysin, aminopeptidase B, THOP, neurotensin-degrading enzyme	22.8	7.5	3.0	-	-	-
bradykinin [42]	ACE, aminopeptidase P, carboxypeptidase N, NEP, DPP-II, DPP-IV,	-	4.2	-	-	3CWW	-
kallidin [42]	ACE, aminopeptidase P, carboxypeptidase N	-	7.3	-	-	-	-
atrial NP (ANP) [43,44]	NEP	-	0.06	-	-	3N57	-
B-type NP (BNP) [44]	NEP, fibroblast activation protein-α	-	weak	-	>1	3N56	-
relaxin [45]	-	-	*0.11**	-	0.054	-	-
relaxin-3 [45]	-	-	*0.36**	-	0.182	-	-
insulin-like peptide 3 [46]	-	0.15	0.055	2.7	-	-	-
β-endorphin (1-31) [42]	NEP	21	13	1.6	-	-	-
growth hormone releasing factor (1−29) [42]	DPP-IV, NEP	23.4	11.9	2.0	-	-	-
pancreastatin (1-49) [42]	-	10.6	41.7	0.25	-	-	-
dynorphin B (1-13) [42]	NEP	15.1	18.3	0.8	-	-	-
A (1-17) [42]	NEP	18.4	37.4	0.5	-	-	-
B (1-9) [42]	NEP	10.3	26.8	0.4	-	-	-
A (1-13) [42]	NEP	3.74	40.6	0.09	-	-	-
A (1-10) [42]	NEP	3.74	39.4	0.09	-	-	-
A (1-8) [42]	NEP	0.66	63.5	0.01	-	-	-
A (1-9) [42]	NEP	0.33	60.5	0.01	-	-	-

Table A3.1 Literature survey of putative IDE substrates identified using *in vitro* assays. (*) K_M estimated based on IC_{50} for inhibition of insulin degradation. Abbreviations: Aβ = amyloid beta; TGF-alpha = transforming growth factor-alpha; IGF = insulin-like growth factor; NP = natriuretic peptide; GRF = gastrin release factor; NEP = neprilysin; PREP = prolyl endopeptidase; ECE = endothelin converting enzyme; THOP = thimet oligopeptidase; ACE = angiotensin-converting enzyme; DPP = dipeptidylpeptidase. Known non-substrates of IDE include glucagon-like peptide 1 (GLP-1), glucose-dependent insulinotropic peptide (GIP), epithelial growth factor 1 (EGF-1) and C-peptide.

Compound	Purity (%)	Formula	[M+H]+ expected	[M+H]+ found	Ion count	Δ (ppm)
1a	87.5	$C_{26}H_{41}N_7O_7$	564.3140	564.3135	17505	-0.886
1b	66.4	$C_{26}H_{41}N_7O_7$	564.3140	564.3141	237080	0.177
2a	84.1	$C_{41}H_{55}N_9O_7$	786.4297	786.4290	523007	-0.890
2b	95.3	$C_{41}H_{55}N_9O_7$	786.4297	786.4272	627894	-3.179
3a	84.3	$C_{46}H_{62}N_6O_8$	827.4702	827.4703	111258	0.121
3b	89.9	$C_{46}H_{62}N_6O_8$	827.4702	827.4703	26066	0.121
4a	93.7	$C_{32}H_{52}N_6O_7$	633.3970	633.3975	1344574	0.789
4b	79.7	$C_{32}H_{52}N_6O_7$	633.3970	633.3967	591800	-0.474
5a	97.1	$C_{41}H_{52}N_6O_7$	741.3970	741.3987	517852	2.293
5b	75.4	$C_{41}H_{52}N_6O_7$	741.3970	741.3967	208124	-0.405
6a	98.3	$C_{40}H_{51}N_7O_8$	758.3872	758.3886	875721	1.846
6b	98.9	$C_{40}H_{51}N_7O_8$	758.3872	758.3878	14154	0.791
7	89.6	$C_{40}H_{53}N_7O_7$	744.4079	744.4079	21411	-0.040
8	94.9	$C_{33}H_{47}N_7O_7$	654.3610	654.3606	413032	-0.611
9	80.1	$C_{37}H_{55}N_7O_7$	710.4236	710.4239	180420	0.422
10	76.9	$C_{39}H_{51}N_7O_7$	730.3923	730.3917	644559	-0.821
11	92.6	$C_{40}H_{45}N_7O_8$	752.3402	752.3399	77585	-0.399
12	96.0	$C_{37}H_{47}N_7O_8$	718.3559	718.3550	145796	-1.253
13	95.7	$C_{37}H_{47}N_7O_9$	718.3559	718.3569	24610	1.392
14	58.2	$C_{34}H_{41}N_7O_8$	676.3089	676.3082	400965	-1.035
15	77.5	$C_{38}H_{48}N_6O_7$	701.3657	701.3671	124002	1.996
16	67.3	$C_{38}H_{48}N_6O_7$	701.3657	701.3683	19335	3.707
17	83.2	$C_{39}H_{50}N_6O_7$	715.3814	715.3808	298036	-0.839
18	92.3	$C_{39}H_{50}N_6O_7$	715.3814	715.3832	701250	2.516
19	37.4	$C_{38}H_{48}N_6O_8$	717.3606	717.3586	128261	-2.788
20	74.5	$C_{37}H_{46}N_6O_7$	687.3501	687.3503	135532	0.291
21	95.6	$C_{40}H_{50}N_6O_9$	759.3712	759.3726	13461	1.844
22	97.9	$C_{40}H_{52}N_6O_8S$	777.3640	777.3634	121542	-0.772
23	88.5	$C_{40}H_{52}N_6O_7Se$	807.3144	807.3139	1497	-0.619
24	52.4	$C_{44}H_{52}N_6O_7$	777.3970	777.3981	15952	1.415
25	63.5	$C_{35}H_{43}N_5O_6$	630.3286	630.3282	136628	-0.635
26	74.4	$C_{40}H_{51}N_7O_8$	758.3872	758.3876	119065	0.527
27	90.9	$C_{39}H_{49}N_7O_8$	744.3715	744.3708	210827	-0.940
29	75.3	$C_{42}H_{55}N_7O_8$	786.4185	786.4178	234651	-0.890
28	94.3	$C_{40}H_{50}N_6O_9$	759.3712	759.3705	89863	-0.922
30	97.5	$C_{54}H_{75}N_9O_{11}S$	1058.5380	1058.5402	23247	2.078
31	81.6	$C_{65}H_{71}N_7O_{15}$	1190.5081	1190.5063	442	-1.512
6bK	96.6	$C_{41}H_{55}N_7O_7$	758.4236	758.4233	501430	-0.396
heavy-6bK	96.4	$C_{35}{}^{13}C_6H_{55}N_5{}^{15}N_2O_7$	766.4378	766.4388	227096	1.305
bisepi-6bK	96.7	$C_{41}H_{55}N_7O_7$	758.4236	758.4232	133009	-0.527
epiA-6bK	96.6	$C_{41}H_{55}N_7O_7$	758.4236	758.4258	1320288	2.901
epiB-6bK	94.9	$C_{41}H_{55}N_7O_7$	758.4236	758.4235	216951	-0.132
epiC-6bK	93.9	$C_{41}H_{55}N_7O_7$	758.4236	758.4235	177995	-0.132
Ii1	94.8	$C_{37}H_{45}N_9O_8$	744.3464	744.3474	662	1.343

Table A3.2 HPLC and high-resolution mass spectrometry analysis of IDE inhibitor analogs.

Mutation	Primers (forward, reverse)
A198Q	ACATGAGAAGAATGTGATGAATGA-dU-CAGTGGAGAC
	ATCATTCATCACATTCTTCTCATG-dU-TCTG
A198T	AGACTCTTTCAATTGGAAAAAGC-dU-ACAGGG
	AGCTTTTTCCAATTGAAAGAGTC-dU-CCAGGTATCATTCATCACATTCTTCTCATGTTC
A198C	AGACTCTTTCAATTGGAAAAAGC-dU-ACAGGG
	AGCTTTTTCCAATTGAAAGAGTC-dU-CCAGCAATCATTCATCACATTCTTCTCATGTTC
A198Y	ACATGAGAAGAATGTGATGAATGA-dU-TACTGGAGAC
	ATCATTCATCACATTCTTCTCATG-dU-TCTG
W199L	ACATGAGAAGAATGTGATGAATGA-dU-GCCTTAAGACTC
	ATCATTCATCACATTCTTCTCATG-dU-TCTG
W199F	AGACTCTTTCAATTGGAAAAAGC-dU-ACAGGG
	AGCTTTTTCCAATTGAAAGAGTC-dU-GAAGGCATCATTCATCACATTCTTCTC
W199Y	AGACTCTTTCAATTGGAAAAAGC-dU-ACAGGG
	AGCTTTTTCCAATTGAAAGAGTC-dU-ATAGGCATCATTCATCACATTCTTCTC
F202L	ACATGAGAAGAATGTGATGAATGA-dU-GCCTGGAGACTCTTGCAATTG
	ATCATTCATCACATTCTTCTCATG-dU-TCTG
F202R	AGACTCTTTCAATTGGAAAAAGC-dU-ACAGGG
	ATGAATGATGCCTGGAGAC-dU-CCGTCAATTGGAAAAAGCTACAGGG
I310R	ACCCATTAAAGATCGTAGGAATC-dU-CTATGTGACATTTCCCATAC
	AGATTCCTACGATCTTTAATGGG-dU-ACTATTTTGTAAAGTTGTTTAAG
Y314F	ACCCATTAAAGATATTAGGAATCTC-dU-TCGTGACATTTCCCATACCTGACCTTC
	AGAGATTCCTAATATCTTTAATGGG-dU-ACTATTTTG
V360Q	AAAGGGCTGGGTTAATACTCT-dU-CAGGGTGGGCAG
	AAGAGTATTAACCCAGCCCTT-dU-GACTTAAG
V360R	AAAGGGCTGGGTTAATACTCT-dU-AGGGGTGGGCAGAAGGAAGGAGC
	AAGAGTATTAACCCAGCCCTT-dU-GACTTAAG
G361Q	AAAGGGCTGGGTTAATACTCT-dU-GTTCAGGGGCAGAAGGAAGGAGCCC
	AAGAGTATTAACCCAGCCCTT-dU-GACTTAAG
G362Q	AAAGGGCTGGGTTAATACTCT-dU-GTTGGTCAGCAGAAGGAAGGAGCCCGAG
	AAGAGTATTAACCCAGCCCTT-dU-GACTTAAG
K364A	AAGGAGCCCGAGGTTTTA-dU-GTTTTTTATC
	ATAAAACCTCGGGCTCCT-dU-CCGCCTGCCCACCAACAAGAGTATT
I374M	ATGTTTTTTATCATGAATGTGGACT-dU-GACCG
	AAGTCCACATTCATGATAAAAAACA-dU-AAAACCTC
I374Q	ATGTTTTTTATCCAGAATGTGGACT-dU-GACCGAGGAAGG
	AAGTCCACATTCTGGATAAAAAACA-dU-AAAACCTCGGGCTCC
A479L	ATGTCCGGGTTCTGATAGTTTCTAAA-dU-CTTTTGAAGGAAAAACTG
	ATTTAGAAACTATCAGAACCCGGACA-dU-TTTCTGGTCTGAG

Table A3.3 Site-directed mutagenesis primers.

Sequencing primers	
Seq_Fw1	GATTAACTTTATTATTAAAAATTAAAGAGG
Seq_Re1	CAACATGTAATAATCCTTCCTCGGTC
Seq_Fw2	GCATGAAGGTCCTGGAAGTCTG
Seq_Re2	AGGAAGGGTTACATCATCCAGAGC
Seq_Fw3	CCATGTACTACCTCCGCTTGC
Seq_Re3	GCAGATCTCGAGCTCGGATC
Seq_Fw4	GCTTATGTGGACCCCTTGCACTG
RT-PCR primers[26,27]	
PEPCK_1_Fw	GAACTGACAGACTCGCCCTATGT
PEPCK_1_Re	GTTGCAGGCCCAGTTGTTG
G6Pase_1_Fw	GTGCAGCTGAACGTCTGTCTGT
G6Pase_1_Re	TCCGGAGGCTGGCATTGT
PEPCK_2_Fw	GGTGTTTACTGGGAAGGCATC
PEPCK_2_Re	CAATAATGGGGCACTGGCTG
G6Pase_2_Fw	CATGGGCGCAGCAGGTGTATACT
G6Pase_2_Re	CAAGGTAGATCCGGGACAGACAG
Tubulin_Fw	CCTGCTCATCAGCAAGATCC
Tubulin_Re	TCTCATCCGTGTTCTCAACC
Beta-actin_ Fw	CATCCGTAAAGACCTCTATGCCAAC
Beta-actin_Re	ATGGAGCCACCGATCCACA

Table A3.4 Sequencing and RT-PCR primers.

Protein	Batch	Relative activity	**Ii1** IC_{50} shift *	**6b** IC_{50} shift	**13** IC_{50} shift	**9** IC_{50} shift	Prediction
hIDE WT	1	1.0	1.0	1.0	1.0	1.0	
hIDE WT	2	1.0	1.0	1.0	1.0	-	
hIDE WT	3	1.0	1.0	1.0	1.0	-	
A198Q	1	1.4	1.0	2.3	-	-	
A198T	2	1.5	1.0	**16**	**13.6**	**-**	scaffold interaction
A198Y	2	2.3	1.1	0.5	-	-	
W199F	2	1.3	1.8	**3.0**	**3.2**	**-**	modest interaction
W199Y	2	0.3	1.4	1.8	-	-	
F202L	1	0.7	1.7	1.7	2.3	-	
F202R	2	0.9	1.1	**3.7**	**4.1**	**-**	modest interaction
I310R	2	3.0	1.9	0.8	0.7	-	
Y314F	2	3.5	1.0	1.2	**-**	**-**	
V360Q	1	2.6	0.9	1.5	-	-	
V360R	3	0.4	0.9	1.4	-	-	
G361Q	3	0.7	0.9	0.8	-	-	
G362Q	3	1.2	0.6	**77**	**1.8**	**-**	cyclohexyl ring interaction
K364A	2	4.3	1.0	**26**	**10.0**	**-**	scaffold interaction
I374M	1	4.3	1.0	0.7	-	-	
A479L	1	0.5	0.9	**>600**	-	**1.0**	*p*-benzoyl interaction

Table A3.5 Site-directed IDE mutants used in the small molecule-enzyme mutant complementation studies. * NB: Ii1 is not known to interact with any of these residues. Significant changes in IC_{50} for Ii1 were presumed to indicate misfolding or other non-inhibitor-specific protein structure changes. For example, W199L displayed an IC_{50} shift of 21-fold for Ii1, 13-fold for 6b, and 17-fold for analog 13.

A3.1 Supplementary References

1 Kleiner, R. E., Dumelin, C. E., Tiu, G. C., Sakurai, K. & Liu, D. R. In vitro selection of a DNA-templated small-molecule library reveals a class of macrocyclic kinase inhibitors. *J Am Chem Soc* **132**, 11779-11791, doi:10.1021/ja104903x (2010).

2 Gartner, Z. J. *et al.* DNA-templated organic synthesis and selection of a library of macrocycles. *Science* **305**, 1601-1605, doi:10.1126/science. 1102629 [pii] (2004).

3 Tse, B. N., Snyder, T. M., Shen, Y. & Liu, D. R. Translation of DNA into a library of 13,000 synthetic small-molecule macrocycles suitable for in vitro selection. *J Am Chem Soc* **130**, 15611-15626, doi:10.1021/ja805649f (2008).

4 Parker, M. G. & Weitzman, P. D. The regulation of Acinetobacter sp. alpha-oxoglutarate dehydrogenase complex. *Biochem J* **130**, 39P (1972).

5 Schenker, P. & Baici, A. Simultaneous interaction of enzymes with two modifiers: reappraisal of kinetic models and new paradigms. *J Theor Biol* **261**, 318-329, doi:10.1016/j.jtbi.2009.07.033 (2009).

6 Kim, Y. G., Lone, A. M., Nolte, W. M. & Saghatelian, A. Peptidomics approach to elucidate the proteolytic regulation of bioactive peptides. *Proc Natl Acad Sci U S A* **109**, 8523-8527, doi:10.1073/pnas.1203195109 (2012).

7 Yang, Q. *et al.* Pharmacological inhibition of BMK1 suppresses tumor growth through promyelocytic leukemia protein. *Cancer cell* **18**, 258-267, doi:10.1016/j.ccr.2010.08.008 (2010).

8 Weisberg, E. *et al.* Beneficial effects of combining a type II ATP competitive inhibitor with an allosteric competitive inhibitor of BCR-ABL for the treatment of imatinib-sensitive and imatinib-resistant CML. *Leukemia* **24**, 1375-1378, doi:10.1038/leu.2010.107 (2010).

9 Wang, L. *et al.* A small molecule modulates Jumonji histone demethylase activity and selectively inhibits cancer growth. *Nature communications* **4**, 2035, doi:10.1038/ncomms3035 (2013).

10 Wang, F. *et al.* Identification of a small molecule with activity against drug-resistant and persistent tuberculosis. *Proc Natl Acad Sci U S A* **110**, E2510-2517, doi:10.1073/pnas.1309171110 (2013).

11 Taipale, M. *et al.* Chaperones as thermodynamic sensors of drug-target interactions reveal kinase inhibitor specificities in living cells. *Nat Biotechnol* **31**, 630-637, doi:10.1038/nbt.2620 (2013).

12 Matzuk, M. M. *et al.* Small-molecule inhibition of BRDT for male contraception. *Cell* **150**, 673-684, doi:10.1016/j.cell.2012.06.045 (2012).

13 Anastasiou, D. *et al.* Pyruvate kinase M2 activators promote tetramer formation and suppress tumorigenesis. *Nat Chem Biol* **8**, 839-847, doi:10.1038/nchembio.1060 (2012).

14 Gottlieb, H. E., Kotlyar, V. & Nudelman, A. NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities. *The Journal of organic chemistry* **62**, 7512-7515 (1997).

15 Georghiou, G., Kleiner, R. E., Pulkoski-Gross, M., Liu, D. R. & Seeliger, M. A. Development and structure-based mechanism of highly specific macrocyclic Src kinase inhibitors from a DNA-template library. *submitted* (2011).

16 Stella, V. J. & He, Q. Cyclodextrins. *Toxicologic pathology* **36**, 30-42, doi:10.1177/0192623307310945 (2008).

17 Shen, Y., Joachimiak, A., Rosner, M. R. & Tang, W. J. Structures of human insulin-degrading enzyme reveal a new substrate recognition mechanism. *Nature* **443**, 870-874, doi:10.1038/nature05143 (2006).

18 Vonrhein, C. *et al.* Data processing and analysis with the autoPROC toolbox. *Acta crystallographica. Section D, Biological crystallography* **67**, 293-302, doi:10.1107/S0907444911007773 (2011).

19 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of applied crystallography* **40**, 658-674, doi:10.1107/s0021889807021206 (2007).

20 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-2132, doi:10.1107/S0907444904019158 (2004).

21 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**, 213-221, doi:doi:10.1107/S0907444909052925 (2010).

22 Lang, P. T. *et al.* DOCK 6: combining techniques to model RNA-small molecule complexes. *Rna* **15**, 1219-1230, doi:10.1261/rna.1563609 (2009).

23 Moustakas, D. T. *et al.* Development and validation of a modular, extensible docking program: DOCK 5. *Journal of computer-aided molecular design* **20**, 601-619, doi:10.1007/s10822-006-9060-4 (2006).

24 Copeland, R. A. *Evaluation of enzyme inhibitors in drug discovery : a guide for medicinal chemists and pharmacologists*. 2nd edn, (Wiley, 2013).

25 Andrikopoulos, S., Blair, A. R., Deluca, N., Fam, B. C. & Proietto, J. Evaluating the glucose tolerance test in mice. *American journal of physiology. Endocrinology and metabolism* **295**, E1323-1332, doi:10.1152/ajpendo.90617.2008 (2008).

26 Llovera, R. E. *et al.* The catalytic domain of insulin-degrading enzyme forms a denaturant-resistant complex with amyloid beta peptide: implications for Alzheimer disease pathogenesis. *J Biol Chem* **283**, 17039-17048, doi:10.1074/jbc.M706316200 (2008).

27 Lee, Y. *et al.* Metabolic manifestations of insulin deficiency do not occur without glucagon action. *Proc Natl Acad Sci U S A* **109**, 14972-14976, doi:10.1073/pnas.1205983109 (2012).

28 Mirsky, I. A. & Broh-Kahn, R. H. The inactivation of insulin by tissue extracts; the distribution and properties of insulin inactivating extracts. *Arch Biochem* **20**, 1-9 (1949).

29 Duckworth, W. C., Bennett, R. G. & Hamel, F. G. Insulin degradation: progress and potential. *Endocr Rev* **19**, 608-624 (1998).

30 Farris, W. *et al.* Insulin-degrading enzyme regulates the levels of insulin, amyloid beta-protein, and the beta-amyloid precursor protein intracellular domain in vivo. *Proc Natl Acad Sci U S A* **100**, 4162-4167, doi:10.1073/pnas.0230450100 0230450100 [pii] (2003).

31 Abdul-Hay, S. O. *et al.* Deletion of insulin-degrading enzyme elicits antipodal, age-dependent effects on glucose and insulin tolerance. *PLoS One* **6**, e20818, doi:10.1371/journal.pone.0020818 PONE-D-10-05674 [pii] (2011).

32 Kurochkin, I. V. & Goto, S. Alzheimer's beta-amyloid peptide specifically interacts with and is degraded by insulin degrading enzyme. *FEBS letters* **345**, 33-37 (1994).

33 Qiu, W. Q. *et al.* Insulin-degrading enzyme regulates extracellular levels of amyloid beta-protein by degradation. *J Biol Chem* **273**, 32730-32738 (1998).

34 Malito, E., Hulse, R. E. & Tang, W. J. Amyloid beta-degrading cryptidases: insulin degrading enzyme, presequence peptidase, and neprilysin. *Cellular and molecular life sciences : CMLS* **65**, 2574-2585, doi:10.1007/s00018-008-8112-4 (2008).

35 Miller, B. C. *et al.* Amyloid-beta peptide levels in brain are inversely correlated with insulysin activity levels in vivo. *Proc Natl Acad Sci U S A* **100**, 6221-6226, doi:10.1073/pnas.1031520100 (2003).

36 Duckworth, W. C. & Kitabchi, A. E. Insulin and glucagon degradation by the same enzyme. *Diabetes* **23**, 536-543 (1974).

37 Shroyer, L. A. & Varandani, P. T. Purification and characterization of a rat liver cytosol neutral thiol peptidase that degrades glucagon, insulin, and isolated insulin A and B chains. *Archives of biochemistry and biophysics* **236**, 205-219 (1985).

38 Bennett, R. G., Duckworth, W. C. & Hamel, F. G. Degradation of amylin by insulin-degrading enzyme. *J Biol Chem* **275**, 36621-36625, doi:10.1074/jbc.M006170200 (2000).

39 Hamel, F. G., Gehm, B. D., Rosner, M. R. & Duckworth, W. C. Identification of the cleavage sites of transforming growth factor alpha by insulin-degrading enzymes. *Biochimica et biophysica acta* **1338**, 207-214 (1997).

40 Misbin, R. I., Almira, E. C., Duckworth, W. C. & Mehl, T. D. Inhibition of insulin degradation by insulin-like growth factors. *Endocrinology* **113**, 1525-1527 (1983).

41 Ciaccio, C. *et al.* Somatostatin: a novel substrate and a modulator of insulin-degrading enzyme activity. *J Mol Biol* **385**, 1556-1567, doi:10.1016/j.jmb.2008.11.025 (2009).

42 Safavi, A., Miller, B. C., Cottam, L. & Hersh, L. B. Identification of gamma-endorphin-generating enzyme as insulin-degrading enzyme. *Biochemistry* **35**, 14318-14325, doi:10.1021/bi960582q (1996).

43 Muller, D., Baumeister, H., Buck, F. & Richter, D. Atrial natriuretic peptide (ANP) is a high-affinity substrate for rat insulin-degrading enzyme. *European journal of biochemistry / FEBS* **202**, 285-292 (1991).

44 Muller, D., Schulze, C., Baumeister, H., Buck, F. & Richter, D. Rat insulin-degrading enzyme: cleavage pattern of the natriuretic peptide hormones ANP, BNP, and CNP revealed by HPLC and mass spectrometry. *Biochemistry* **31**, 11138-11143 (1992).

45 Bennett, R. G., Heimann, D. G. & Hamel, F. G. Degradation of relaxin family peptides by insulin-degrading enzyme. *Annals of the New York Academy of Sciences* **1160**, 38-41, doi:10.1111/j.1749-6632.2008.03782.x (2009).

46 Zhang, W. J., Luo, X. & Guo, Z. Y. In vitro degradation of insulin-like peptide 3 by insulin-degrading enzyme. *The protein journal* **29**, 93-98, doi:10.1007/s10930-009-9226-8 (2010).